FUNDAMENTAL IMPACTS OF DELAY AND DEADLINE ON COMMUNICATIONS OVER WIRELESS NETWORKS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sina Lashgari

February 2016

FUNDAMENTAL IMPACTS OF DELAY AND DEADLINE ON COMMUNICATIONS OVER WIRELESS NETWORKS

Sina Lashgari, Ph.D.

Cornell University 2016

Consumer demand for wireless networks has experienced an exponential increase in recent years, and such trend is expected to continue in the coming years. Therefore, there is a pressing need for more effective communication schemes to be able to cope with the explosive growth in the wireless demand. An important resource which can be very useful in developing more effective communications schemes is channel state information which is available at the transmitters (CSIT). CSIT is usually provided to the transmitters via feedback channels from the receivers. Timely and accurate CSIT has been shown to provide enormous theoretical gains in terms of communication rate. However, in most realistic scenarios, instantaneous and perfect CSIT is not feasible due to physical constraints on the feedback channel. Furthermore, the CSIT supplied by different receivers can often times be of different quality in terms of timeliness and accuracy, which results in networks with heterogeneous/hybrid CSIT.

In this dissertation, we provide new tools and techniques to better understand and analyze the fundamental limits of wireless networks under practical CSIT constraints. In particular, we develop mathematical tools that capture the impact of various types of CSIT on the received signal dimensions at different receivers in a wireless network. We also show how the developed tools are used to solve a broad spectrum of problems in network information theory, from interference networks such as X-Channel with delayed CSIT, 3-user Interference

Channel with delayed CSIT, Interference Channel with limited transmitter co-operation and delayed CSIT, and 2-by-k multiple-input single-output broadcast channel (MISO BC) with delayed CSIT, to various problems in information-theoretic security, and finally, MISO BC under heterogeneous CSIT. The developed tools presented in this dissertation provide new insights on the impact of CSIT on the dynamics of wireless networks.

Beside the significance of quality of CSIT in wireless networks, quality of service (QoS) requirements by the traffic also play an important role in the design of better communication schemes. The majority of increasing traffic volume over wireless networks is expected to be video, which is in most cases delay-sensitive; i.e., packets should be delivered by a certain time, otherwise they will not be useful to the user. Therefore, we study fundamental limits of communicating delay-sensitive traffic over heterogeneous wireless networks, which are emerging structures in modern wireless networks. We provide approximate characterization of the timely throughput capacity of such heterogeneous wireless networks, and develop near-optimal algorithms.

BIOGRAPHICAL SKETCH

Sina Lashgari started his B.Sc. studies in the Electrical Engineering department at Sharif University of Technology in 2006 after being ranked in the top 0.01% of students in the nationwide university entrance exam. His field of concentration was Communications Engineering. During his B.Sc. studies Sina received the Outstanding Academic Achievements Award from the president of Sharif University, and National Elites Foundation fellowship. He finished his B.Sc. in 2010, and joined Cornell University in August 2010 to study Ph.D. at the Electrical and Computer Engineering department. He was the recipient of Jacobs Fellowship at Cornell University, and is a member of Foundation of Information Engineering (FoIE) at Cornell.

During his Ph.D., Sina has developed novel mathematical techniques that help better understand and analyze the dynamics of wireless networks under practical constraints. He has also collaborated with Intel, Cisco, and Verizon on addressing the emerging challenges in the next generation of wireless networks. During his internship at the R&D division of Qualcomm in summer 2013, he led a project on developing new algorithms for Device-to-Device communications in large-scale wireless networks. Moreover, during his internship at the Securities division of Goldman Sachs in summer 2015, he conducted quantitative research that led to the development of new equities trading strategies. His main research interests lie in the areas of Information Theory, Statistics, and Machine Learning.

To my parents, Parviz and Parvin,

whose love and support cannot be described by words,

and to my love, Pegah,

who is the joy of my life.

ACKNOWLEDGEMENTS

First, I would like to sincerely thank my advisor, Prof. Salman Avestimehr, for all his help and support throughout my PhD studies. Salman is extremely committed to the professional and academic development of his students. He always emphasized the importance of tackling hard and big problems in a systematic way, and I am sure learning his approach of tackling challenging problems will be very valuable to me later on in my career. In addition, he helped me a lot on my journey to become an independent researcher. His passion for research is always contagious to people who work with him; and his ambition for tackling challenging research problems and breaking the barriers has always been inspiring. Finally, he helped me a lot in improving my presentation and writing skills, for which I will be always grateful.

Also I would like to thank Prof. Tsuhan Chen, my co-advisor, who has been incredibly supportive, and whose comments and advice have been very helpful to me during my PhD program. Professor Chen has always been very kind and approachable, and I certainly benefited a lot from his advice. In addition, I would like to thank Prof. Adrian Lewis, who has also been on my PhD committee. I had the great opportunity of taking his optimization course which I enjoyed very much; and I am lucky to have him on my committee.

During my PhD studies, I also had the great opportunity to work with several brilliant researchers, including Prof. Changho Suh from KAIST, Dr. Javad Abdoli from Huawei, and Prof. Ravi Tandon from University of Arizona. I learned a lot from them and enjoyed our discussions very much, and I would like to cordially thank them for the opportunity of collaborating with them, and for their insightful comments. Moreover, I had the chance to conduct research in my last year of B.Sc. at Sharif University under supervision of Prof. Farid

Ashtiani. I learned a lot from Prof. Ashtiani, got inspired by his character and work ethics, and I am always thankful for his help and support during my last year as an undergraduate student.

Furthermore, I had the opportunity to do two internships during my Ph.D. studies, which helped enrich my experience and broaden my vision. I would like to thank Dr. Shreeshankar Bodas and Dr. Saurabh Tavildar who were my mentors at Qualcomm for all their help and support during the internship. I would also like to thank my supervisors at Goldman Sachs, Mr. Jesse Cohen and Ms. Kelly Brennan, for investing so much time and effort and making the internship an amazing learning opportunity for me.

I am also very thankful to Ashkan Hajjam, Navid Naderi, and Hamid Pezeshkian who have made my graduate life much more enjoyable. Ashkan has always been like a brother to me; and I feel lucky to have amazing friends such as Hamid and Navid. They were definitely missed after leaving Cornell.

I would like to thank Scott Coldren, who has always been incredibly helpful during my time at Cornell. Scott is one of the few people in Cornell who goes way out of his way to help students. Also, many thanks to Daniel Richter for all his help and support through these years. Daniel really cares about the people around him and his presence clearly made Rhodes Hall a better place. I enjoyed our friendly chats during lunch time very much.

Finally, I would like to thank my family, Pegah, Parviz, Maliheh, and Kiana, who have always shown unwavering and unconditional support for my every endeavor, and in every struggle of my life. They are the ones who have cherished my happy moments and my accomplishments, and have always been by my side during tough times. I certainly feel blessed to have such a wonderful family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Consumer demand for data services over wireless networks has increased dramatically in recent years, fueled both by the success of online video streaming and popularity of smartphones and tablet devices. This confluence of trends is expected to continue and lead to several fold increase in traffic over wireless networks in the next few years [1]. As a result, one of the most pressing challenges in wireless networks is to find better communication schemes that are scalable and can serve the future needs of wireless traffic while considering practical constraints of such networks.

An important and useful resource for developing more effective communication schemes is the channel state information which is available at the transmitters (CSIT). Timely and accurate CSIT can be utilized by the transmitters to employ smarter communication schemes which improve the communication rate in wireless networks. Nevertheless, the common procedure for obtaining CSIT is to send training symbols (or pilots) at the transmitters, and then estimate the channels at the receivers and feed the estimates back to the transmitters. As a result of this feedback mechanism, it is not always reasonable to assume that CSIT is perfect and instantaneous. For instance, CSIT may be outdated due to the fast fading nature of the channels or slow feedback mechanism, it can be noisy (imperfect), or not available at all.

Hence, there are two important questions: how can delayed CSIT be used in wireless networks to develop better communication schemes? and what are the fundamental limits of communications over wireless networks under delayed CSIT? As it turns out, there is still a clear lack of understanding regarding networks with delayed CSIT, to the extent that even for simple network configurations such as X-channel and three-user interference channel, the fundamental limits of communications are unknown.

In fact, although two main converse techniques (genie-aided channel enhancement [65] and statistical equivalence of channel outputs [87]) have been developed in the literature for networks with delayed CSIT, the existing techniques in network information theory fail to address a broad range of networks under delayed CSIT. Hence, there is a need for new tools and techniques that allow for better analysis of the dynamics of networks under delayed CSIT. In Chapter 2, we develop novel tools that allow for better analysis of some of the most fundamental wireless networks in network information theory under delayed CSIT.

On the other hand, there are networks in which not all receivers supply CSIT. For instance, consider the wiretap channel, which is one of the canonical settings in the information-theoretic study of secrecy in wireless networks. It consists of a transmitter that wishes to communicate a secret message to a legitimate receiver in the presence of eavesdropper(s) that should not decode the secret message. In such network it is not reasonable to assume that eavesdropper(s) would cooperate with the transmitter by supplying CSIT. However, assuming the legitimate receiver supplies delayed CSIT, secure degrees of freedom (SDoF) of such network has only been solved for the case where the eavesdropper(s) also provides delayed CSIT [95]. Hence, an interesting problem is to consider wiretap channel where the eavesdropper(s) does not supply any CSIT, and the legitimate receiver only supplies delayed CSIT. We study this problem and develop new techniques to solve the problem for two different network configurations, which are described in Section 1.2.2 and Chapter 3.

The problem of different receivers providing different types of CSIT (due to their distinct feedback channels) can be studied in a broader setting, where each receiver can supply CSIT of a different quality. This results in communication scenarios with *heterogeneous* (or *hybrid*) CSIT. As a result, there have also been several works on studying the impact of heterogeneous (or hybrid) CSIT on the capacity of wireless networks, where the CSIT with respect to each receiver can now be either instantaneous/perfect, delayed, or not available. However, studying networks under the assumption of heterogeneous CSIT becomes quite challenging, to the extent that for k-user multiple-input single-output broadcast channel (MISO BC), the degrees of freedom (DoF) is only characterized for $k = 2$ [20, 81]; and beyond the 2-user network configuration even the DoF is unknown and the problem remains widely open. In Chapter 4 we study the problem of broadcast channel with hybrid CSIT beyond two users, and we develop a new technique which is used to settle the case of 3-user MISO BC, and provide new results on the general k-user MISO BC with hybrid CSIT.

Aside from quality of CSIT, which has significant impacts on wireless communications, the quality of service (QoS) requirements by wireless traffic play an important role in communications over wireless networks as well. Consumer demand for data services over wireless networks has increased dramatically in recent years; and the majority of such demand is video. As a result, the majority of wireless traffic is time-sensitive, i.e., packets should be delivered by a certain time, otherwise they will not be useful anymore.

With the evolution of wireless networks towards heterogeneous architectures, including wireless relays and femtocells, and growing number of smart devices that can connect to several wireless technologies (e.g. 3G and WiFi), it is promising that the opportunistic utilization of heterogeneous networks (where available) can be one of the key solutions to help cope with the phenomenal growth of video demand over wireless networks. This motivates two fundamental questions: first, how much is the ultimate capacity gain from opportunistic utilization of network heterogeneity for delay-sensitive traffic? and second, what are the optimal policies that exploit network heterogeneity for delivery of delay-sensitive traffic? In Chapter 5 we study this problem in detail and provide new results.

1.2 Prior Works

In this section we describe the problems considered in each chapter, and then provide the relevant existing works in the academic literature.

1.2.1 Fundamental Limits of Interference Management with Delayed CSIT

In Chapter 2 we study fundamental limits of interference management with delayed CSIT, and in particular, we study four problems: X-channel with delayed CSIT, three-user interference channel with delayed CSIT, interference channel with limited transmitter cooperation and delayed CSIT, and 2-by-k multiple-input single-output broadcast channel (MISO BC) with delayed CSIT. We briefly describe each problem here, and present the relevant existing results in the academic literature.

The *X*-channel is a canonical setting for the information-theoretic study of interference management in wireless networks. This channel consists of two transmitters causing interference at two receivers, and each transmitter aims to communicate intended messages to both receivers (Fig. 2.1). On the other hand, three-user interference channel consists of three transmitters having distinct messages for three receivers, where each transmitter causes interference at two unintended receivers (Fig. 2.2). The question is: how can the transmitters optimally manage the interference and communicate their messages to the receivers? This problem has been studied extensively in the literature and various interference management techniques have been proposed. In particular, in [35] it is shown that, quite surprisingly, one can significantly improve upon conventional interference management schemes (e.g., orthogonalization) and achieve 4/3 degrees of freedom (DoF) by using *interference alignment* (IA) [64, 16].

However, in order to perfectly align the interference, the transmitters need to accurately know the *current* state of the channels, which is practically very challenging and may even be impossible (due to, for example, high mobility). In the context of broadcast channel, Maddah-Ali and Tse in [65] have shown that delayed CSIT can still be very useful. In particular, for the multi-antenna broadcast channel with delayed CSIT, they developed an innovative transmission strategy that utilizes the past received signals to create signals of common interest to multiple receivers, hence significantly improving DoF by broadcast-ing them to the receivers.

Subsequently in [66, 86, 6, 25], the impact of delayed CSIT has been explored for a variety of interference networks in which transmit antennas are now distributed at different locations. Unlike multi-antenna broadcast channels, in networks with distributed transmitters, it may not be possible for a transmitter to reconstruct previously received signals, since it may include other transmitters' signals that are not accessible to that transmitter. Hence, although interference alignment has happened in the past receptions, it may not be possible to construct the aligned interference locally at a transmitter and broadcast it to the receivers. Interestingly, even in this setting, delayed CSIT has shown to still provide DoF gains (see e.g., [66, 86, 6, 7, 25]). In particular, for the X-channel, Ghasemi-Motahari-Khandani in [25] developed a scheme that achieves DoF of $\frac{6}{5}$ with delayed CSIT, which is strictly larger than its DoF with no-CSIT (i.e., 1 DoF).

There have also been several converse techniques developed in the literature for networks with delayed CSIT. For the MISO broadcast channel with delayed CSIT, Maddah-Ali and Tse [65] have provided an upper bound based on the genie-aided bounding technique. This technique essentially consists of two steps. First, signals of a set of receivers are given to other set of receivers such that the enhanced network becomes a physically degraded broadcast channel. Using the fact that feedback cannot increase capacity for physically degraded broadcast channels [23], we can then take the non-feedback upper bound as that of the original feedback channel. This technique has also been used in [24] in the context of broadcast erasure channels with feedback. Furthermore, for time correlated MISO broadcast channel with delayed CSIT, a converse has been proposed in [93], which is based on extremal inequality [62]. Moreover, for MIMO interference channel with delayed CSIT, a converse has been proposed in [86], which utilizes the fact that for delayed CSIT, the signals received at different receivers in a time slot are statistically equivalent; therefore, the entropy of received signals at different receivers in a certain time slot are equal when conditioned on past received signals at any specific receiver. However, the existing techniques fail to characterize the DoF for X-channel with delayed CSIT. We present our contributions in this regard in Section 1.3.1.

Another problem that we focus on in Chapter 2 is two-user interference channel with delayed CSIT, which is the fundamental setting for the information-theoretic study of interference management in wireless networks. This channel consists of two transmitters causing interference at two receivers, and each transmitter aims to communicate a message to its intended receiver (see Fig. 2.3 for the configuration).

For two-user interference channel where the channels are time-varying, one can show that time-sharing (i.e. TDMA) is optimal in terms of degrees of freedom (DoF); and therefore, DoF is 1. On the other hand, when the two transmitters can fully cooperate and share their messages, the network turns into two-user multiple-input single-output broadcast channel (MISO BC). For the MISO BC with instantaneous CSIT, one can show that DoF=2. However, in order to perfectly align the interference, the transmitter needs to accurately know the *current* state of the channels, which is practically very challenging and may even be impossible. It was shown in [65] that DoF $= \frac{4}{3}$ for two-user MISO broadcast channel with delayed CSIT.

However, at a higher level, one can view the two-user interference channel and two-user MISO BC as two extreme cases of transmitter cooperation, where the former corresponds to no cooperation, while the latter corresponds to full cooperation, in which each transmit antenna has access to both messages.

The problem of transmitter cooperation for interference channel has been considered in prior works including [67, 32, 89, 88, 37, 3]. In particular, in [67, 32, 89] a model for partial transmitter cooperation in two-user interference channel is considered in which one transmitter shares its entire message with the other transmitter. In particular, [67] studies the capacity of discrete memoryless (DMC) interference channel, while [32, 89] consider the Gaussian interference channel with time-invariant channels. However, the form of cooperation considered in [67, 32, 89] is specific to certain networks, and as it can be seen later in Chapter 2, is subsumed by our model as a special case.

Moreover, in [88, 37] interference channel with conferencing transmitters is considered; that is, cooperative links are orthogonal to each other as well as to the links in the interference channel. Nevertheless, the type of cooperation considered in our work is different in the sense that instead of considering a channel with certain capacity between the transmitters, we abstract the message sharing mechanism by assuming that the transmitters have partial access to one another's messages. On the other hand, there have been works such as [3] which focus on full-duplex cooperation between the transmitters. However, such capability might not be present on the transmitter side in many network scenarios. We present our contribution on interference channel with limited cooperation and delayed CSIT in Section 1.3.1.

Another problem that we will study in Chapter 2 is k-user MISO BC, where the transmitter has a distinct message to communicate to each receiver. The transmitter is equipped with 2 antennas, while each receiver is equipped with a single antenna, as depicted in Fig. 2.7. The transmitter has access to the delayed CSIT; hence, the problem is called the $2 \times k$ MISO BC with delayed CSIT.

For this problem there have been lower and upper bounds developed on its DoF in [65]. However, the bounds are only tight for the special cases of $k = 2, 3$, and beyond that, there is a clear gap between the existing lower and upper bounds.

1.2.2 Information-Theoretic Security with Practical CSIT Constraints

In Chapter 3 we study two problems: (i) blind MIMOME wiretap channel with delayed CSIT (Fig. 3.1), and blind cooperative SISO wiretap channel with delayed CSIT (Fig. 3.4). Wiretap channel consists of a transmitter that wishes to communicate a secret message to a legitimate receiver in the presence of eavesdropper(s) that should not decode the confidential message. There has been a large amount of work on this problem, and its secrecy capacity has been determined in several configurations (e.g., [90, 18, 61, 60]). In particular, the secrecy capacity of the Gaussian wiretap channel is characterized in [60], and it is known that if the channel to the legitimate receiver is "less noisy" than the channel to the eavesdropper, then a positive rate of secret communication is achievable.

However, the secrecy capacity of the Gaussian wiretap channel does not scale with the available transmit power, i.e., the secure degrees of freedom

(SDoF) of Gaussian wiretap channel is zero. This has motivated the utilization of helping jammers and multi-antenna transmitters in networks to increase the achievable SDoF (e.g. [28, 40, 41, 91, 92, 80, 11, 98, 97, 27, 39]). In particular, it has been shown in [91] that the SDoF of wiretap channel with a helping jammer (i.e. cooperative jamming) in a wireless setting in which the channels remain constant is $\frac{1}{2}$. This work has also been extended in [92] to the setting with no eavesdropper CSIT (i.e., blind cooperative jamming). However, the above results rely on assuming that channels are constant, and do not change over time.

Secure communication over networks with time-varying channels (i.e. ergodic channels) has been considered in some prior works in the literature [43, 26, 95, 42, 71]. In particular, in [43], achievability results for SDoF of K-user interference channel with instantaneous CSIT were presented. Moreover, SDoF of wireless X-networks has been studied in [26]. Nevertheless, the results in these works heavily rely on the assumption that the transmitters have perfect and instantaneous CSIT.

Therefore, there have been follow up works that focus on studying SDoF for settings in which only delayed CSIT is available. In particular, in [95] Yang et al. have considered the Gaussian MIMO wiretap channel with delayed CSIT; and they have characterized the SDoF of such network for arbitrary number of antennas. However, they assume that the eavesdropper supplies the transmitter with perfect delayed CSIT, which in most scenarios is not a realistic assumption. For the case where no eavesdropper CSIT is available, [95] provides lower bounds on the SDoF which only match their respective upper bounds for specific network configurations, and the SDoF is in general unknown.

1.2.3 MISO Broadcast Channel with Heterogeneous CSIT

As we mentioned, there has been a growing interest in studying the impact of CSIT on the capacity of wireless networks, especially the broadcast channel. In particular, it was shown in [65] that for k-user MISO BC with delayed CSIT, DoF $= \frac{k}{1+\frac{1}{2}+\ldots+\frac{1}{k}}$. This work was followed by several other works which studied other network configurations under the assumption of delayed CSIT, including interference channel [6, 86, 84, 50], X-channel [25, 52], multi-hop networks [4], and other variations of delayed CSIT [93].

Most of these prior works assume that the entire network state information is obtained with delay. However, in a large network, one can expect various types of CSIT to be available at the transmitters with respect to different receivers. As a result, there have also been several works on studying the impact of heterogeneous (or hybrid) CSIT on the capacity of wireless networks, where the CSIT with respect to each receiver can now be either instantaneous/perfect (P), delayed (D), or not available (N) [66, 82, 70, 81, 72, 69, 9]. However, studying networks under the assumption of heterogeneous CSIT becomes quite challenging, to the extent that only the DoF for 2-user MISO BC is characterized [20, 81]; and beyond the 2-user network configuration even the DoF is unknown and the problem remains widely open.

1.2.4 Timely Throughput of Heterogeneous Wireless Networks

In Chapter 5 we study the downlink of a heterogeneous wireless network with N Access Points (AP's) and M clients. We assume that each AP is using a distinct frequency band, and all AP's are connected to each other through a Backhaul Network, with error free links. We model the wireless channels as packet erasure channels (see Fig. 5.1(a)). Time is slotted and time-slots are grouped to form intervals of length τ. For each interval every client has packets to receive and the AP's have to decide on a scheduling policy to deliver the packets. If a packet is not delivered by the end of that interval, it gets dropped by the AP's. Total timely throughput, T^3, is defined as the long-term average number of successful deliveries in the network. Our objective is then to find the maximum achievable T^3, which we denote by C_{T^3}, over all possible scheduling policies.

Although there are classical results [24], [73] on scheduling clients over time-varying channels and characterizing the average delay of service, in recent years there has been increasing research on serving delay-sensitive traffic over wireless networks [2, 8, 74, 76].

However, the most related work to Chapter 5 is the work of Hou et al. in [34] in 2009, in which they have proposed a framework for jointly addressing delay, delivery ratio, and channel reliability. For a network with one AP and N clients, the timely throughput region for the set of N clients has been fully characterized in [34]; and the work has been extended to variable-bit-rate applications in [29], and time-varying channels and rate adaptation in [30]. Although in [34]- [30] they provide tractable analytical results and low-complexity scheduling policies, the analyses are done for only one AP. In fact, timely throughput region for $N = 1$ can be shown to be a scaled version of a polymatroid [33]. However, once we move beyond $N = 1$, the timely throughput region loses its polymatroidal structure which makes the problem much more challenging. Chapter 5 aims to extend the results to the case of general number of AP's.

1.3 Overview of Contributions

In this section we provide an overview of main contributions in each chapter.

1.3.1 Fundamental Limits of Interference Management with Delayed CSIT

Given that the only upper bound on the DoF of X-channel with delayed CSIT is the one with instantaneous CSIT (i.e., $\frac{4}{3}$ DoF), it still remains open whether $\frac{6}{5}$ is the fundamental limit on the DoF of X-channel with delayed CSIT, or whether there are more efficient interference management techniques.

In Chapter 2 we show that the DoF of the Gaussian X- channel with delayed CSIT is indeed $\frac{6}{5}$, under the assumption that only linear encoding schemes are employed at the transmitters. Under this constraint, only a linear combination of information symbols are allowed to be transmitted at each time. Linear schemes are of significant practical interests due to their low complexity; and in fact, the majority of DoF-optimal schemes developed so far for networks with delayed CSIT are linear (e.g., [65, 66, 86, 6, 25]).

The key part of the converse is the development of a general lemma, namely "Rank Ratio Inequality", that bounds the maximum ratio of the dimensions of received linear-subspaces (at the two receivers) that are created by *distributed* transmitters with delayed CSIT. More specifically, we show that if two distributed transmitters with delayed CSIT employ linear strategies, the ratio of the dimensions of the received signals cannot exceed $\frac{3}{2}$. With instantaneous

CSIT, this ratio can be as large as 2, and with no CSIT, this ratio is always 1. As a result, this lemma captures the fundamental impact of delayed CSIT on the dimension of received subspaces. Also, in the case of two centralized transmitters (e.g., multi-antenna BC), this ratio can be as large as 2, therefore Rank Ratio Inequality also captures the impact of *distributed transmitters* on the dimension of received subspaces.

We also demonstrate how our lemma can be applied to any arbitrary network, in which a receiver decodes its desired message in the presence of two interferers. As an example, we apply the lemma to the three-user interference channel with delayed CSIT and derive a new upper bound of $\frac{9}{7}$ on its linear DoF. This is the first upper bound that captures the impact of delayed CSIT on the degrees of freedom of this network.

We then consider the two-user Gaussian interference channel with fading channels and delayed CSIT in Chapter 2. We first present a model to capture and quantify the amount of cooperation between the transmitters. In this model we denote the fraction of shared messages that are intended for $\mathrm{Rx}_1, \mathrm{Rx}_2$ by ρ_1, ρ_2, respectively, and then, characterize the degrees of freedom (DoF) region as a function of ρ_1, ρ_2. As a result, the two-user interference channel and two-user multiple-input single-output broadcast channel (MISO BC) become special cases of no cooperation ($\rho_1 = \rho_2 = 0$) and full cooperation ($\rho_1 = \rho_2 = 1$) in our framework. Moreover, our result indicates that the maximum benefit of cooperation from the DoF perspective is achieved by sharing only half of the messages between the transmitters.

The proof of achievability is based on the observation that in order to achieve the DoF tuple $(\frac{2}{3}, \frac{2}{3})$ for the two-user interference channel with delayed CSIT, it is sufficient for each transmitter to have access to half of the symbols available to the other transmitter. The converse however, is based on a new lemma which characterizes the maximum amount of interference alignment at a receiver as a function of amount of message sharing. In particular, Lemma 6 implies that once d_1 DoF is delivered to Rx_1, and at most a fraction ρ_1 of W_1 is shared with Tx_2, then the interference at Rx_2 occupies at least $(1 - \rho_1)d_1$ dimensions. Using this lemma we can bound the maximum DoF that can be communicated to each receiver, and prove the converse. The lemma is analogous to the Rank Ratio Inequality. However, the assumptions of the two lemmas are different: in this lemma (Lemma 6) we assume each receiver is able to decode a certain amount of DoF, and that messages are partially shared among transmitters; while in the Rank Ratio Inequality no decodability assumption is made; and messages are not shared.

Finally, in Chapter 2 we study the 2-by-k MISO BC with delayed CSIT. We first provide a new achievable scheme for 2×4 MISO BC with delayed CSIT, which improves the state-of-the-art in [65]. In particular, our scheme achieves $\frac{14}{9}$ DoF via a four-phase scheme which employs a more efficient interference alignment using delayed CSIT. We then provide a generalization of our achievable scheme for $2 \times k$ MISO BC with delayed CSIT.

1.3.2 Information-Theoretic Security with Practical CSIT Constraints

In Chapter 3 we first consider blind MIMOME wiretap channel with delayed CSIT and completely characterize its SDoF for all antenna configurations. In particular, we improve the state-of-the-art achievable schemes in [95]; and we provide tight upper bounds on the SDoF.

In our proposed achievable scheme the transmitter transmits artificial noise symbols in order to perform two tasks simultaneously: first, the artificial noise signals span the entire received signal space at the eavesdroppers to completely drown the confidential message in noise at the eavesdroppers. Second, artificial noise signals are aligned into a smaller linear subspace at the legitimate receiver in order to occupy less signal dimensions and leave some room for the confidential message to be decoded.[1] Our achievable scheme performs these two tasks by utilizing the delayed CSIT provided by the legitimate receiver in a two-phase transmission scheme. For settings in which the legitimate receiver has less antennas than an eavesdropper, our proposed achievable scheme allows for more efficient artificial noise alignment at the eavesdroppers by spending less time slots for generating artificial noise equations for retransmission, hence improving the achieved SDoF.

The converse proof is based on 4 main lemmas. Each lemma presents an inequality which provides a lower bound on the received signal dimension at a receiver which supplies a certain type of CSIT. These inequalities provide the essential tools for analyzing the received signal dimensions at different receivers in blind MIMOME wiretap channel with delayed CSIT. In particular, Least Alignment Lemma, states that if two receivers in a network have the same number of antennas and one of the receivers supplies no CSIT, the least amount of alignment will occur at that receiver, meaning that transmit signals will occupy the maximal signal dimensions at that receiver.

[1]Artificial noise Alignment was introduced in [41] to mask the confidential message in the artificial noise at the undesired receivers.

Moreover, Lemma 8 and Lemma 10 provide lower bounds on the received signal dimensions at receivers which supply delayed and no CSIT, respectively. Finally, Lemma 11 provides a lower bound on the received signal dimensions at a collection of receivers, where some receivers supply no CSIT.

We then consider the problem of blind cooperative SISO wiretap channel with delayed CSIT, where the secure communication is aided via a distributed jammer, where a jammer is a transmitter that does not necessarily have access to the confidential message, but can help jam the confidential message at the eavesdropper(s). All nodes in the network have a single antenna. We characterize the linear SDoF of such network, which is SDoF when transmitters are restricted to use linear encoding schemes. Converse proof is derived by utilizing the Rank Ratio Inequality (Lemma 1) along with Least Alignment Lemma (Lemma 12).

1.3.3 MISO Broadcast Channel with Heterogeneous CSIT

To make progress on the k-user MISO BC with hybrid CSIT beyond 2 users, we focus on characterizing the degrees of freedom when restricted to linear schemes (also called LDoF). We first study the case of $k = 3$, and fully characterize the LDoF for all 3^3 possible hybrid CSIT configurations. The result is obtained by developing a general outer bound on the LDoF region, and a matching achievable scheme for each of the CSIT configurations.

The outer bound, is based on three main ingredients. The first ingredient is a novel lemma, called *Interference Decomposition Bound*. It essentially lower bounds the interference dimension at a receiver with delayed CSIT by the average dimension of its constituents, thereby decomposing the interference into its individual components. As a result of Interference Decomposition Bound, we can then focus on analyzing the dimension of constituents of interference at receivers which supply delayed CSIT, in order to derive an upper bound on LDoF. Proof of Interference Decomposition Bound is based on temporal analysis of dimensions of transmit signals at different receivers, leading to necessary conditions on the increments of such dimensions using the delayed CSIT constraint.

The second main ingredient of the converse proof is *MIMO Rank Ratio Inequality for Broadcast Channel*, which provides a lower bound on the dimension of interference components at receivers supplying delayed CSIT. Its equivalent version for general encoding schemes has been presented in Chapter 3 to prove the converse for blind MIMOME wiretap channel with delayed CSIT. In particular, the bound states that if the transmitter employs linear precoding schemes, the dimension of each interference component at a single-antenna receiver supplying delayed CSIT is at least half of the dimension of the corresponding signal at any other single-antenna receiver.

Finally, the third ingredient of the converse, is *Least Alignment Lemma*, a variation of which is presented in Chapter 3 to prove the converse for blind cooperative SISO wiretap channel with delayed CSIT. Using the three main ingredients we develop a converse proof which characterizes the LDoF region for all 3^3 possible hybrid CSIT configurations of the 3-user MISO BC.

We next extend the key proof ingredients of the converse for 3-user MISO BC to the general k-user setting. By extending the converse tools to the general k-user setting, we provide a new outer bound on the linear DoF region of the general k-user MISO BC with arbitrary hybrid CSIT configuration. We demonstrate that our new outer bound leads to an approximate linear sum-DoF characterization to within an additive gap of 0.5 for networks with more number of receivers supplying instantaneous CSIT than delayed CSIT; and the approximation gap decays exponentially with the increase in number of receivers supplying instantaneous CSIT. Furthermore, by using the outer bound and providing a new multi-phase achievable scheme, we present the exact characterization of linear sum-DoF for networks in which only one receiver supplies delayed CSIT.

1.3.4 Timely Throughput of Heterogeneous Wireless Networks

The challenge in characterizing C_{T^3} is that for each interval, even the number of different ways of assigning packets to AP's is N^M, which grows exponentially in the number of clients (M). For $N = 1$, timely throughput region is a scaled version of a polymatroid [33]. However, once we move beyond $N = 1$, the timely throughput region loses its polymatroidal structure which makes the problem much more challenging. To overcome the challenge, we propose a *deterministic relaxation* of the problem, which is based on converting the problem to a network with deterministic delays for each link. As we will show in Chapter 5, the relaxed problem can be viewed as an assignment problem in which each AP turns into a bin with certain capacity and each packet turns into an object which has different sizes at different bins. The relaxed problem is then to maximize the total number of objects that can be packed in the bins, denoted by C_{det}.

Our main contribution is two-fold. First, we prove that the gap between the solutions to the original problem (C_{T^3}) and its relaxed version (C_{det}) is at most

$2\sqrt{N(C_{\text{det}} + \frac{N}{4})}$. Since N is typically very small (in most cases between 2-4), the above result indicates that C_{det} is asymptotically equal to C_{T^3} as $C_{\mathsf{T}^3} \to \infty$. Furthermore, our numerical results demonstrate that the gap is in most cases much smaller than the worst-case gap that we prove analytically. Therefore, instead of solving our main maximization problem we can solve its relaxed version, and still get a value which is very close to the optimum. Second, we prove that the relaxed problem can be approximated in polynomial-time (with additive gap of N) using a simple LP rounding method. This approximation is appealing as N is usually limited and negligible compared to C_{det}. As a result, the solution to the relaxed problem provides a scheduling policy that provably achieves a T^3 that is within additive gap $N + 2\sqrt{N(C_{\mathsf{T}^3} - \frac{3N}{4})}$ of C_{T^3} for $C_{\mathsf{T}^3} > \frac{7N}{4}$.

We also consider several extensions of the problem, including extension to time-varying channels and real-time traffic, where at the beginning of each interval clients have requests for variable number of packets. We show that the aforementioned results hold in these two extensions, too. Moreover, we provide similar results for the case where different flows have different priorities (different weights). In addition, we extend the model to allow for online scheduling policies, where AP's are coordinated, and a packet might be transmitted by arbitrary number of AP's. Finally, we consider an extension to account for fading, multiple simultaneous transmissions by AP's and multiple simultaneous receptions by clients, and rate adaptation.

CHAPTER 2

FUNDAMENTAL LIMITS OF INTERFERENCE MANAGEMENT WITH DELAYED CSIT

2.1 Overview

With the scarcity of the spectrum, and the explosive growth in wireless traffic volume, users are often forced to communicate in the same time and frequency, causing interference at one another. As a result, there is a clear need for effective interference management techniques. On the other hand, it has been shown that channel state information at the transmitters (CSIT) can improve interference management (see e.g. [16, 65]). However, for time-varying channels CSIT is usually outdated by the time it gets to the transmitters. Therefore, it is important to study interference management in light of delayed CSIT. In this chapter we study the impacts of delayed CSIT on the fundamental limits of interference management in wireless interference networks. In particular, we study many important interference networks such as X-channel, interference channels with/without transmitter cooperation, and 2-by-k multiple-input single-output broadcast channel (MISO BC) with delayed CSIT.

We first focus on X-channel, which is a canonical setting for the information-theoretic study of interference management in wireless networks.[1] It consists of two transmitters causing interference at two receivers, and each transmitter aims to communicate intended messages to both receivers. We show that the Degrees of Freedom (DoF) of the Gaussian X- channel with delayed CSIT is indeed $\frac{6}{5}$, under the assumption that only linear encoding schemes are employed

[1]The results presented in this chapter have been presented in part in [52, 51, 50].

at the transmitters. Under this constraint, only a linear combination of information symbols are allowed to be transmitted at each time. The key part of the converse is the development of a general lemma, namely "Rank Ratio Inequality", that bounds the maximum ratio of the dimensions of received linear-subspaces (at the two receivers) that are created by *distributed* transmitters with delayed CSIT.

We also demonstrate how Rank Ratio Inequality can be applied to any arbitrary network in which a receiver decodes its desired message in the presence of two interferers. As an example, we apply the lemma to the three-user interference channel with delayed CSIT and derive a new upper bound of $\frac{9}{7}$ on its linear DoF.

We then focus on another fundamental setting for the information-theoretic study of interference management: two-user interference channel. This channel consists of two transmitters causing interference at two receivers, and each transmitter aims to communicate a message to its intended receiver (see Fig. 2.3). We consider the two-user Gaussian interference channel with fading channels and delayed CSIT. We propose a model for capturing transmitter cooperation, and completely characterize the DoF of two-user interference channel with delayed CSIT and partial transmitter cooperation. The converse proof is based on a key lemma which characterizes the maximum amount of interference alignment at a non-intended receiver as a function of the amount of message sharing.

Finally, we consider the problem of k-user multiple-input single-output broadcast channel (MISO BC), where there are k single-antenna receivers and a 2-antenna transmitter, as depicted in Fig. 2.7. The transmitter has access to the delayed CSIT; hence, the problem is called the $2 \times k$ MISO BC with delayed

CSIT. We first provide a new achievable scheme for 2×4 MISO BC with delayed CSIT, which improves the state-of-the-art scheme presented in [65]. We then generalize our achievable scheme for $2 \times k$ MISO BC with delayed CSIT.

2.2 X-Channel with Delayed CSIT

In this section we study the impact of delayed channel state information at the transmitters (CSIT) on interference management in the context of X-channel, which is a canonical setting for the information-theoretic study of interference management in wireless networks. We first present the system model and main result, which is characterization of Degrees of Freedom (DoF) under linear schemes. We then prove the result by developing a general lemma, namely "Rank Ratio Inequality", that bounds the maximum ratio of the dimensions of received linear-subspaces (at the two receivers) that are created by *distributed* transmitters with delayed CSIT.

2.2.1 System Model & Main Results

Throughout this Section, we use small letters for scalars, arrowed letters (e.g. $\vec{x}$) for vectors, capital letters for matrices, and a calligraphic font for sets. Furthermore, we use bold letters for random entities, and non-bold letters for deterministic values (e.g., realizations of random variables).

We consider the Gaussian X-channel depicted in Fig. 2.1. It consists of two transmitters and two receivers, and each transmitter has a separate message for each of the receivers. Each node is equipped with a single antenna.

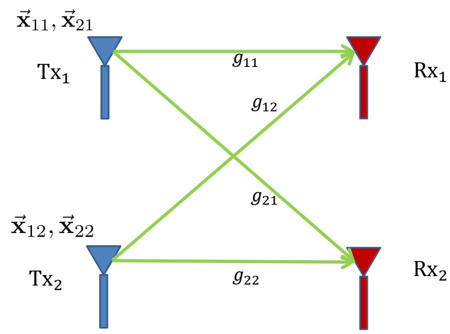


Figure 2.1: Network configuration for X-channel. There are two transmitters and two receivers, where each transmitter has a message for each receiver. We assume time-varying channels, with delayed CSIT.

The received signal at Rx_k ($k \in \{1, 2\}$) at time t is given by

$$\mathbf{y}_k(t) = \mathbf{g}_{k1}(t)\mathbf{x}_1(t) + \mathbf{g}_{k2}(t)\mathbf{x}_2(t) + \mathbf{z}_k(t), \tag{2.1}$$

where $\mathbf{x}_j(t)$ is the transmit signal of Tx_j; $\mathbf{g}_{kj}(t) \in \mathbb{C}$ indicates a channel from Tx_j to Rx_k; and $\mathbf{z}_k(t) \sim \mathcal{CN}(0, 1)$. The channel coefficients of $\mathbf{g}_{kj}(t)$'s are i.i.d across time and users, and they are drawn from a continuous distribution. We denote by $\mathcal{G}(t)$ the set of all four channel coefficients at time t. In addition, we denote by $\mathcal{G}^n$ the set of all channel coefficients from time 1 to n, i.e.,

$$\mathcal{G}^n = \{\mathbf{g}_{kj}(t) : k, j \in \{1, 2\}, t = 1, \ldots, n\}.$$

Denoting the vector of transmit signals for Tx_j in a block of length n by $\vec{\mathbf{x}}_j^n$, each transmitter Tx_j obeys an average power constraint, $\frac{1}{n}E\{\|\vec{\mathbf{x}}_j^n\|^2\} \leq P$. We assume delayed channel state information at the transmitters (CSIT). In other words, at time t, only the states of the past $\mathcal{G}^{t-1}$ are known to the transmitters. Furthermore, we assume that receivers have instantaneous CSIT, meaning that at time t, $\mathcal{G}^t$ is known to all receivers.

We restrict ourselves to linear coding strategies as defined in [13], in which DoF simply represents the dimension of the linear subspace of transmitted signals. More specifically, consider a communication scheme with block length n, in which transmitter Tx_j wishes to transmit a vector $\vec{\mathbf{x}}_{kj} \in \mathbb{C}^{m_{kj}(n)}$ of $m_{kj}(n) \in \mathbb{N}$ information symbols to Rx_k ($j,k \in \{1,2\}$). These information symbols are then modulated with precoding vectors $\vec{\mathbf{v}}_{kj}(t) \in \mathbb{C}^{m_{kj}(n)}$ at times $t = 1, 2, \ldots, n$. Note that the precoding vector $\vec{\mathbf{v}}_{kj}(t)$ depends only upon the outcome of $\mathcal{G}^{t-1}$ due to the delayed CSIT constraint:

$$\vec{v}_{kj}(t) = f^{(n)}_{k,j,t}\left(\mathcal{G}^{t-1}\right). \tag{2.2}$$

Based on this linear precoding, Tx_j will then send $\mathbf{x}_j(t) = \vec{\mathbf{v}}_{1j}(t)^\top \vec{\mathbf{x}}_{1j} + \vec{\mathbf{v}}_{2j}(t)^\top \vec{\mathbf{x}}_{2j}$ at time t. We denote by $\mathbf{V}^n_{kj} \in \mathbb{C}^{n \times m_{kj}(n)}$ the overall precoding matrix of Tx_j for Rx_k, such that the t-th row of $\mathbf{V}^n_{kj}$ is $\vec{\mathbf{v}}_{kj}(t)^\top$). In addition, we denote the precoding functions used by Tx_j by $f^{(n)}_j = \{f^{(n)}_{1,j,t}, f^{(n)}_{2,j,t}\}^n_{t=1}$, $j = 1, 2$.

Based on the above setting, the received signal at Rx_k ($k \in \{1,2\}$) after the n time steps of the communication will be

$$\vec{\mathbf{y}}^n_k = \mathbf{G}^n_{k1}(\mathbf{V}^n_{11}\vec{\mathbf{x}}_{11} + \mathbf{V}^n_{21}\vec{\mathbf{x}}_{21}) + \mathbf{G}^n_{k2}(\mathbf{V}^n_{12}\vec{\mathbf{x}}_{12} + \mathbf{V}^n_{22}\vec{\mathbf{x}}_{22}) + \vec{\mathbf{z}}^n_k, \tag{2.3}$$

where $\mathbf{G}^n_{kj}$ is the $n \times n$ diagonal matrix whose t-th element on the diagonal is $\mathbf{g}_{kj}(t)$. [2] Now, consider the decoding of $\vec{\mathbf{x}}_{kj}$ at Rx_k (i.e., the $m_{kj}(n)$ information symbols of Tx_j for Rx_k). The corresponding interference subspace at Rx_k will be

$$\mathcal{I}_{kj} = \text{colspan}\left(\mathbf{G}^n_{kj}\mathbf{V}^n_{k'j}\right) \cup \text{colspan}\left(\mathbf{G}^n_{kj'}\mathbf{V}^n_{kj'}\right) \cup \text{colspan}\left(\mathbf{G}^n_{kj'}\mathbf{V}^n_{k'j'}\right),$$

where $j' = 3 - j, k' = 3 - k$, and colspan(.) of a matrix corresponds to the subspace that is spanned by its columns. For instance, $\mathcal{I}_{11} = \text{colspan}(\mathbf{G}^n_{11}\mathbf{V}^n_{21}) \cup$

[2] For $j,k \in \{1,2\}$, we define $\mathbf{G}^0_{kj}\mathbf{V}^0_{kj} \triangleq 0_{1\times m_{kj}(n)}$; therefore, for instance, we have $\text{rank}\left[\mathbf{G}^0_{k1}\mathbf{V}^0_{k1} \quad \mathbf{G}^0_{k2}\mathbf{V}^0_{k2}\right] = 0$, $k \in \{1,2\}$.

colspan($\mathbf{G}_{12}^n \mathbf{V}_{12}^n$) ∪ colspan($\mathbf{G}_{12}^n \mathbf{V}_{22}^n$). Let $\mathcal{I}_{kj}^c \subseteq \mathbb{C}^n$ denote the subspace orthogonal to $\mathcal{I}_{kj}$. Then, in the regime of asymptotically high transmit powers (i.e., ignoring the noise), the decodability of information symbols from Tx_j at Rx_k corresponds to the constraints that the image of colspan($\mathbf{G}_{kj}^n \mathbf{V}_{kj}^n$) on $\mathcal{I}_{kj}^c$ has dimension $m_{kj}(n)$:

$$\dim\left(\mathrm{Proj}_{\mathcal{I}_{kj}^c} \mathrm{colspan}\left(\mathbf{G}_{kj}^n \mathbf{V}_{kj}^n\right)\right) = \dim\left(\mathrm{colspan}\left(\mathbf{V}_{kj}^n\right)\right) = m_{kj}(n). \tag{2.4}$$

Based on this setting, we now define the sum linear degrees of freedom of the X-channel.

Definition 1. *Four-tuple $(d_{11}, d_{12}, d_{21}, d_{22})$ degrees of freedom are linearly achievable if there exists a sequence $\{f_1^{(n)}, f_2^{(n)}\}_{n=1}^{\infty}$ such that for each n and the choice of $(m_{11}(n), m_{12}(n), m_{21}(n), m_{22}(n))$, $(\mathbf{V}_{11}^n, \mathbf{V}_{12}^n, \mathbf{V}_{21}^n, \mathbf{V}_{22}^n)$ satisfy the decodability condition of (4.9) with probability 1, and $\forall (j,k)$,*

$$d_{kj} = \lim_{n\to\infty} \frac{m_{kj}(n)}{n}. \tag{2.5}$$

We also define the linear degrees of freedom region $\mathcal{D}$ as the closure of the set of all achievable 4-tuples $(d_{11}, d_{12}, d_{21}, d_{22})$. Furthermore, the sum linear degrees of freedom ($\mathrm{DoF}_{\text{L-sum}}$) is then defined as follows:

$$\mathrm{DoF}_{\text{L-sum}} = \max \sum_{k,j\in\{1,2\}} d_{kj}, \quad \text{s.t. } (d_{11}, d_{12}, d_{21}, d_{22}) \in \mathcal{D}. \tag{2.6}$$

In case transmitters have instantaneous CSIT, it was shown in [64, 15] that the sum degrees of freedom is $\frac{4}{3}$. The achievability uses interference alignment that enables us to deliver four symbols over three timeslots. On the other hand, in the non-CSIT case, one can readily see that the received signals at the two receivers are statistically identical and therefore the DoF collapses to 1, which is that of the multiple access channel. For the case of delayed CSIT, Ghasemi-Motahari-Khandani in [25] develops a new scheme that achieves the sum DoF of $\frac{6}{5}$.

Our main result in this section is the following theorem, proved in Section 2.2.2, which states that $\frac{6}{5}$ is the maximum DoF that can be achieved using linear encoding schemes.

Theorem 1. *For the X-channel with delayed CSIT,*

$$\mathrm{DoF}_{\text{L-sum}} = \frac{6}{5}. \tag{2.7}$$

Our converse proof builds upon the following key lemma, which is proved in Section 2.2.2.

Lemma 1. (Rank Ratio Inequality) *For any linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_{11}^n, \mathbf{V}_{12}^n$ as defined in (2.2),*

$$\operatorname{rank}\left[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n\right] \overset{a.s.}{\leq} \frac{3}{2}\operatorname{rank}\left[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n\right]. \tag{2.8}$$

Remark 1. *Note that this lemma holds for any arbitrary network (or sub-network) with two transmitters and two receivers. It does not require any specific decodability assumption at receivers. The inequality of* (2.8) *says that the ratio of the ranks of received beamforming matrices at* Rx_1 *and* Rx_2 *is at most* $\frac{3}{2}$. *For the case of having instantaneous CSIT, one can show that the ratio of* $\operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n]$ *to* $\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]$ *can be up to 2.*[3] *Hence, Lemma 1 characterizes the impact of delayed CSIT on the maximum ratio of the ranks of received beamforming matrices.*

[3]To see this, consider the following two-timeslot scheme. In time 1, $\mathrm{Tx}_1, \mathrm{Tx}_2$ send $\mathbf{x}_1, \mathbf{x}_2$ respectively. Rx_2 then gets $\mathbf{g}_{21}(1)\mathbf{x}_1 + \mathbf{g}_{22}(1)\mathbf{x}_2$. In time 2, $\mathrm{Tx}_1, \mathrm{Tx}_2$ send $\frac{\mathbf{g}_{21}(1)}{\mathbf{g}_{21}(2)}\mathbf{x}_1$, $\frac{\mathbf{g}_{22}(1)}{\mathbf{g}_{22}(2)}\mathbf{x}_2$ respectively. Rx_2 then gets the same equation as the one received in time 1. On the other hand, Rx_1 gets a new equation almost surely. Therefore, the rank of the received signal at Rx_1 can be twice that of Rx_2. Also one can readily show that the two is the maximum that can be achieved.

2.2.2 Proof of Theorem 1

In this section we will prove Theorem 1.

Achievability

As mentioned in the previous section, the achievability is provided in [25], and utilizes a linear encoding scheme to achieve $\frac{6}{5}$. Here we review the scheme to illustrate how beamforming vectors are chosen. We set $n = 5, m_{11}(n) = 2, m_{12}(n) = 1, m_{21}(n) = 1, m_{22}(n) = 2$. Let the information symbols of the transmitters be denoted by

$$\vec{\mathbf{x}}_{11} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \vec{\mathbf{x}}_{12} = \begin{bmatrix} b_1 \end{bmatrix},$$
$$\vec{\mathbf{x}}_{21} = \begin{bmatrix} c_1 \end{bmatrix}, \quad \vec{\mathbf{x}}_{22} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}. \tag{2.9}$$

In $t = 1$, Tx_1 sends a_1, and Tx_2 sends b_1, which corresponds to choosing the following beamforming vectors at the transmitters

$$\vec{v}_{11} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{v}_{12} = \begin{bmatrix} 1 \end{bmatrix}, \vec{v}_{21} = \begin{bmatrix} 0 \end{bmatrix}, \vec{v}_{22} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In $t = 2$, Tx_1 sends a_2, and Tx_2 sends b_1, which corresponds to choosing the following beamforming vectors at the transmitters

$$\vec{v}_{11} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{v}_{12} = \begin{bmatrix} 1 \end{bmatrix}, \vec{v}_{21} = \begin{bmatrix} 0 \end{bmatrix}, \vec{v}_{22} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Therefore, by the end of $t = 2$, Rx_2 can cancel b_1 from its received signals to recover an equation only involving a_1 and a_2, denoted by $\vec{\mathbf{m}}_1^\top \vec{\mathbf{x}}_{11}$. It is easy to see that, if this equation is delivered to Rx_1, it can decode all of its desired information symbols (i.e., $\vec{\mathbf{x}}_{11}$ and $\vec{\mathbf{x}}_{12}$). Hence, it is an equation of interest to Rx_1 that is known at Rx_2, and can be created by Tx_1.

A similar schemes is applied in the next two time steps. More specifically, in $t = 3$, Tx_1 sends c_1, and Tx_2 sends d_1, which corresponds to choosing the following beamforming vectors at the transmitters

$$\vec{v}_{11} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \vec{v}_{12} = \begin{bmatrix} 0 \end{bmatrix}, \vec{v}_{21} = \begin{bmatrix} 1 \end{bmatrix}, \vec{v}_{22} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

In $t = 4$, Tx_1 sends c_1, and Tx_2 sends d_2, which corresponds to choosing the following beamforming vectors at the transmitters

$$\vec{v}_{11} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \vec{v}_{12} = \begin{bmatrix} 0 \end{bmatrix}, \vec{v}_{21} = \begin{bmatrix} 1 \end{bmatrix}, \vec{v}_{22} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Therefore, by the end of $t = 4$, Rx_1 can cancel c_1 from its received signals to recover an equation only involving d_1 and d_2, denoted by $\vec{\mathbf{m}}_2^\top \vec{\mathbf{x}}_{22}$. Again, it is easy to see that, if this equation is delivered to Rx_2, it can decode all of its desired information symbols (i.e., $\vec{\mathbf{x}}_{21}$ and $\vec{\mathbf{x}}_{22}$). Hence, it is an equation of interest to Rx_2 that is known at Rx_1, and can be created by Tx_2.[4]

Now, in $t = 5$, Tx_1 sends $\vec{\mathbf{m}}_1^\top \vec{\mathbf{x}}_{11}$, and Tx_2 sends $\vec{\mathbf{m}}_2^\top \vec{\mathbf{x}}_{22}$. Since each of these transmit signals is already known at one of the receivers, after this transmission, Rx_1 will recover $\vec{\mathbf{m}}_1^\top \vec{\mathbf{x}}_{11}$ and Rx_2 will recover $\vec{\mathbf{m}}_2^\top \vec{\mathbf{x}}_{22}$. Therefore, all information symbols are delivered to their corresponding receivers, achieving sum DoF of $\frac{6}{5}$.

[4]One can check that $\vec{\mathbf{m}}_1 = [\mathbf{g}_{22}(2)\mathbf{g}_{21}(1) \quad - \mathbf{g}_{22}(1)\mathbf{g}_{21}(2)]^\top$, and $\vec{\mathbf{m}}_2 = [\mathbf{g}_{12}(3)\mathbf{g}_{11}(4) \quad - \mathbf{g}_{11}(3)\mathbf{g}_{12}(4)]^\top$.

Converse

We will now prove the converse, which is the main contribution of this section. As mentioned in Section 5.2, the key idea behind the converse is Lemma 1, which we restate below (proof of Lemma 1 is provided in Section 2.2.2).

Lemma 1. (Rank Ratio Inequality) *For any linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_{11}^n, \mathbf{V}_{12}^n$ *as defined in (2.2),*

$$\mathrm{rank}\,[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n] \overset{a.s.}{\leq} \frac{3}{2}\mathrm{rank}\,[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]\,. \tag{2.10}$$

To prove the converse we also need the following three lemmas. The following lemma states the sub-modularity property of rank of matrices (see [63] for more details).

Lemma 2. (Sub-modularity of rank) *Consider a matrix* $A^{m\times n} \in \mathbb{C}^{m\times n}$*. Let* A_I*,* $I \subseteq \{1, 2, \ldots, n\}$ *denote the sub-matrix of* A *created by those columns in* A *which have their indices in* I*. Then, for any* $I_1, I_2 \subseteq \{1, 2, \ldots, n\}$ *we have*

$$\mathrm{rank}[A_{I_1}] + \mathrm{rank}[A_{I_2}] \geq \mathrm{rank}[A_{I_1\cap I_2}] + \mathrm{rank}[A_{I_1\cup I_2}]. \tag{2.11}$$

The following lemma is helpful in providing an equivalent condition for decodability of messages in (4.9), whose proof is based on basic linear algebra and omitted.

Lemma 3. *For two matrices* A, B *of the same row size,*

$$\dim(\mathrm{Proj}_{\mathrm{colspan}(B)^c}\mathrm{colspan}(A)) = \mathrm{rank}[A \quad B] - \mathrm{rank}[B], \tag{2.12}$$

where $\mathrm{Proj}_{\mathrm{colspan}(B)^c}\mathrm{colspan}(A)$ *is the orthogonal projection of column span of* A *on the orthogonal complement of the column span of* B*.*

Finally, the following lemma, whose proof is based on the sub-modularity of the rank function (Lemma 2), will be useful later in the converse proof.

Lemma 4. *Suppose that for four matrices A, B, C, D with the same number of rows,*

$$
\begin{aligned}
\operatorname{rank}[A] + \operatorname{rank}[B \quad C \quad D] &= \operatorname{rank}[A \quad B \quad C \quad D], \\
\operatorname{rank}[B] + \operatorname{rank}[A \quad C \quad D] &= \operatorname{rank}[A \quad B \quad C \quad D].
\end{aligned} \tag{2.13}
$$

Then,

$$
\operatorname{rank}[A] + \operatorname{rank}[B] + \operatorname{rank}[C \quad D] = \operatorname{rank}[A \quad B \quad C \quad D].
$$

Proof. Note that $\operatorname{rank}[A] + \operatorname{rank}[B] + \operatorname{rank}[C \quad D] \geq \operatorname{rank}[A \quad B \quad C \quad D]$. Hence, in order to prove Lemma 4, we only need to prove the inequality in the other direction. Now, according to the assumptions in the Lemma, and using sub-modularity of the rank (Lemma 2), we have

$$
\begin{aligned}
\operatorname{rank}[A] + \operatorname{rank}[B] &\overset{(2.13)}{=} \operatorname{rank}[A \quad B \quad C \quad D] - \operatorname{rank}[B \quad C \quad D] \\
&\quad + \operatorname{rank}[A \quad B \quad C \quad D] - \operatorname{rank}[A \quad C \quad D] \\
&\overset{\text{(sub-modularity)}}{\leq} \operatorname{rank}[A \quad B \quad C \quad D] - \operatorname{rank}[B \quad C \quad D] \\
&\quad + \operatorname{rank}[B \quad C \quad D] - \operatorname{rank}[C \quad D] \\
&= \operatorname{rank}[A \quad B \quad C \quad D] - \operatorname{rank}[C \quad D].
\end{aligned} \tag{2.14}
$$

$\square$

We are now ready to prove the converse. In particular, we prove the following two inequalities:

$$
(d_{11} + d_{12}) + \frac{3}{2}(d_{21} + d_{22}) \leq \frac{3}{2} \tag{2.15}
$$

$$
\frac{3}{2}(d_{11} + d_{12}) + (d_{21} + d_{22}) \leq \frac{3}{2}. \tag{2.16}
$$

The desired result follows from summing the above two inequalities. By symmetry, we only need to prove (2.15). Suppose $(d_{11}, d_{12}, d_{21}, d_{22}) \in \mathcal{D}$, i.e., there exists a sequence $\{f_1^{(n)}, f_2^{(n)}\}_{n=1}^{\infty}$ resulting in linearly achieving $\{m_{11}(n), m_{12}(n), m_{21}(n), m_{22}(n)\}_{n=1}^{\infty}$ with probability 1, and $d_{kj} = \lim_{n\to\infty} \frac{m_{kj}(n)}{n}$. First, note that

$$\dim\left(\text{colspan}(\mathbf{V}_{kj}^n)\right) \overset{a.s.}{=} \dim\left(\text{colspan}(\mathbf{G}_{kj}^n\mathbf{V}_{kj}^n)\right), \tag{2.17}$$

due to the continuous distribution of $\mathbf{g}_{kj}(t)$ for any t. Therefore, by (2.17) and Lemma 3, we conclude that if (4.9) occurs with probability 1, then for $j, k \in \{1, 2\}$ and $j' = 3 - j, k' = 3 - k$,

$$\begin{aligned}
&\text{rank}[\mathbf{G}_{kj}^n\mathbf{V}_{k'j}^n \quad \mathbf{G}_{kj'}^n\mathbf{V}_{kj'}^n \quad \mathbf{G}_{kj'}^n\mathbf{V}_{k'j'}^n] + \text{rank}[\mathbf{G}_{kj}^n\mathbf{V}_{kj}^n] \\
&\overset{a.s.}{=} \text{rank}[\mathbf{G}_{k1}^n\mathbf{V}_{k1}^n \quad \mathbf{G}_{k2}^n\mathbf{V}_{k2}^n \quad \mathbf{G}_{k1}^n\mathbf{V}_{k'1}^n \quad \mathbf{G}_{k2}^n\mathbf{V}_{k'2}^n].
\end{aligned} \tag{2.18}$$

Thus, we consider (2.18) as the equivalent decodability condition, which consists of the following four equations:

$$\begin{aligned}
&\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n] + \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \\
&\overset{a.s.}{=} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n], &(2.19)\\
&\text{rank}[\mathbf{G}_{12}^n\mathbf{V}_{12}^n] + \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \\
&\overset{a.s.}{=} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n], &(2.20)\\
&\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{21}^n] + \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] \\
&\overset{a.s.}{=} \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n], &(2.21)\\
&\text{rank}[\mathbf{G}_{22}^n\mathbf{V}_{22}^n] + \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n] \\
&\overset{a.s.}{=} \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n]. &(2.22)
\end{aligned}$$

Hence, by (2.19), (2.20), and Lemma 4,

$$\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n] + \text{rank}[\mathbf{G}_{12}^n\mathbf{V}_{12}^n] \overset{a.s.}{=} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n]$$

$$- \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n]. \tag{2.23}$$

In addition, by (2.21), (2.22), and Lemma 4,

$$\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{21}^n] + \operatorname{rank}[\mathbf{G}_{22}^n\mathbf{V}_{22}^n] \overset{a.s.}{=} \operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n]$$
$$- \operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]. \tag{2.24}$$

Therefore, we have

$$\begin{aligned}
& m_{11}(n) + m_{12}(n) + \frac{3}{2}(m_{21}(n) + m_{22}(n)) \\
& \overset{a.s.}{=} \operatorname{rank}[\mathbf{V}_{11}^n] + \operatorname{rank}[\mathbf{V}_{12}^n] + \frac{3}{2}(\operatorname{rank}[\mathbf{V}_{21}^n] + \operatorname{rank}[\mathbf{V}_{22}^n]) \\
& \overset{a.s.}{=} \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n] + \operatorname{rank}[\mathbf{G}_{12}^n\mathbf{V}_{12}^n] + \frac{3}{2}(\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{21}^n] + \operatorname{rank}[\mathbf{G}_{22}^n\mathbf{V}_{22}^n]) \\
& \overset{(2.23),\,(2.24)}{\underset{a.s.}{=}} \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] - \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{22}^n] \\
& \quad + \frac{3}{2}(\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] - \operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n]) \\
& \overset{(a)}{\leq} \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{12}^n\mathbf{V}_{12}^n] \\
& \quad + \frac{3}{2}\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] - \frac{3}{2}\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n] \\
& \overset{(\text{Lemma 1})}{\underset{a.s.}{\leq}} \frac{3}{2}\operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_{11}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{12}^n \quad \mathbf{G}_{21}^n\mathbf{V}_{21}^n \quad \mathbf{G}_{22}^n\mathbf{V}_{22}^n] \leq \frac{3}{2}n,
\end{aligned} \tag{2.25}$$

where (a) follows from the fact that $\operatorname{rank}[\mathbf{A} \quad \mathbf{B}] \leq \operatorname{rank}[\mathbf{A}] + \operatorname{rank}[\mathbf{B}]$. Therefore, by dividing both sides of the inequality in (2.25) by n, and letting $n \to \infty$ we get

$$d_{11} + d_{12} + \frac{3}{2}(d_{21} + d_{22}) \leq \frac{3}{2}. \tag{2.26}$$

Hence, the proof of converse for Theorem 1 is complete. ■

We will next prove Lemma 1.

Proof of Lemma 1

Let us fix $n \in \mathbb{N}$, and consider a fixed linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_{11}^n, \mathbf{V}_{12}^n$ as defined in (2.2). For notational simplicity in the proof, we denote $\mathbf{V}_{11}^n$ by $\mathbf{V}_1^n$, and $\mathbf{V}_{12}^n$ by $\mathbf{V}_2^n$. We first state some definitions.

Definition 2. *Consider a fixed linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_1^n \triangleq \mathbf{V}_{11}^n, \mathbf{V}_2^n \triangleq \mathbf{V}_{12}^n$*. Define the random set* $\boldsymbol{\mathcal{T}}_{\{f_1^{(n)}, f_2^{(n)}\}}(\boldsymbol{\mathcal{G}}^n)$ *with its alphabet being the power set of* $\{1, 2, \ldots, n\}$ *as follows. For any realization of channels* $\boldsymbol{\mathcal{G}}^n = \mathcal{G}^n$*, which results in* $\mathbf{G}_{21}^n = G_{21}^n, \mathbf{G}_{22}^n = G_{22}^n, \mathbf{G}_{11}^n = G_{11}^n, \mathbf{G}_{12}^n = G_{12}^n$*, and* $\mathbf{V}_1^n = V_1^n, \mathbf{V}_2^n = V_2^n$*, we define*

$$\mathcal{T}_{\{f_1^{(n)}, f_2^{(n)}\}}(\mathcal{G}^n) \triangleq \{t | [\vec{v}_1(t)^\top \quad \vec{0}_{1\times m_2(n)}], [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(t)^\top] \in rowspan[G_{21}^{t-1} V_1^{t-1} \quad G_{22}^{t-1} V_2^{t-1}]\}.$$

In words, $\boldsymbol{\mathcal{T}}_{\{f_1^{(n)}, f_2^{(n)}\}}(\boldsymbol{\mathcal{G}}^n)$ represents the set of random timeslots (random due to the randomness in channels), where the beamforming vectors transmitted by the two transmitters are already individually recoverable by Rx_2 using its received beamforming vectors in the previous timeslots. Since the code $\{f_1^{(n)}, f_2^{(n)}\}$ is fixed in the proof, for notational simplicity from now on we denote $\boldsymbol{\mathcal{T}}_{\{f_1^{(n)}, f_2^{(n)}\}}(\boldsymbol{\mathcal{G}}^n)$ by $\boldsymbol{\mathcal{T}}$.

Definition 3. *Consider a fixed linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_1^n \triangleq \mathbf{V}_{11}^n, \mathbf{V}_2^n \triangleq \mathbf{V}_{12}^n$*. Define random variables* $\mathbf{r}_1(\boldsymbol{\mathcal{G}}^n), \mathbf{r}_2(\boldsymbol{\mathcal{G}}^n)$ *in* $\{1, \ldots, n\}$ *as follows. For any realization of channels* $\boldsymbol{\mathcal{G}}^n = \mathcal{G}^n$*, which results in* $\mathbf{G}_{21}^n = G_{21}^n, \mathbf{G}_{22}^n = G_{22}^n, \mathbf{G}_{11}^n = G_{11}^n, \mathbf{G}_{12}^n = G_{12}^n$*, and* $\mathbf{V}_1^n = V_1^n, \mathbf{V}_2^n = V_2^n$*, define*

$$r_i(\mathcal{G}^n) \triangleq \dim\left(\text{span}(\mathcal{E}_i(\mathcal{G}^n))\right), \quad i = 1, 2,$$

where

$$\mathcal{E}_1(\mathcal{G}^n) \triangleq \{\vec{s}_{m_1(n)\times 1} | \quad \exists \vec{l}_{n\times 1} \quad s.t. [\vec{s}^\top \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^\top [G_{21}^n V_1^n \quad G_{22}^n V_2^n]\},$$

$$\mathcal{E}_2(\mathcal{G}^n) \triangleq \{\vec{s}_{m_2(n)\times 1} | \quad \exists \vec{l}_{n\times 1} \quad s.t. [\vec{0}_{1\times m_1(n)} \quad \vec{s}^{\top}] = \vec{l}^{\top}[G_{21}^n V_1^n \quad G_{22}^n V_2^n]\}.$$

In words, $\mathbf{r}_1(\mathcal{G}^n)$ can be interpreted as the number of linearly independent equations that Rx_2 can recover from its received signal, which only involve symbols of Tx_1. Hereafter, we denote $\mathbf{r}_1(\mathcal{G}^n), \mathbf{r}_2(\mathcal{G}^n)$ simply by $\mathbf{r}_1, \mathbf{r}_2$.

We will now state the following lemma, proved in Appendix A.1, which is the key to proving Lemma 1.

Lemma 5. *For any linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_1^n \overset{\Delta}{=} \mathbf{V}_{11}^n, \mathbf{V}_2^n \overset{\Delta}{=} \mathbf{V}_{12}^n$ *defined in (2.2),*

- $\text{rank}[\mathbf{G}_{11}^n \mathbf{V}_1^n \quad \mathbf{G}_{12}^n \mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n] \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{11}^{\mathcal{T}} \mathbf{V}_1^{\mathcal{T}} \quad \mathbf{G}_{12}^{\mathcal{T}} \mathbf{V}_2^{\mathcal{T}}]$
- $\text{rank}[\mathbf{V}_j^{\mathcal{T}}] \leq \mathbf{r}_j, \quad j = 1, 2$
- $\mathbf{r}_j \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n] - \text{rank}[\mathbf{V}_{3-j}^n], \quad j = 1, 2$

where $\mathcal{T}$ *is defined in Definition 2,* $\mathbf{V}_i^{\mathcal{T}}$ *represents the random sub-matrix of* $\mathbf{V}_i^n$ *derived by keeping rows whose indices are in* $\mathcal{T}$*, and* $\mathbf{r}_1, \mathbf{r}_2$ *are defined in Definition 3.*

Remark 2. *Note that the first inequality in the above lemma intuitively implies that, in order to bound the difference of the dimensions of received linear subspaces at the two receivers, we only needs to focus on the timeslots in which* Rx_2 *already knows* <u>*both*</u> *of the individual transmit equations.*

We are now ready to prove Lemma 1. We will first use Lemma 5 to find an upper bound on the difference between $\text{rank}[\mathbf{G}_{11}^n \mathbf{V}_1^n \quad \mathbf{G}_{12}^n \mathbf{V}_2^n]$ and $\text{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n]$.

$$\text{rank}[\mathbf{G}_{11}^n \mathbf{V}_1^n \quad \mathbf{G}_{12}^n \mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n] \overset{\substack{(\text{Lemma 5}) \\ a.s.}}{\leq} \text{rank}[\mathbf{G}_{11}^{\mathcal{T}} \mathbf{V}_1^{\mathcal{T}} \quad \mathbf{G}_{12}^{\mathcal{T}} \mathbf{V}_2^{\mathcal{T}}]$$

$$
\begin{aligned}
&\leq \text{rank}[\mathbf{G}_{11}^{\mathcal{T}}\mathbf{V}_1^{\mathcal{T}}] + \text{rank}[\mathbf{G}_{12}^{\mathcal{T}}\mathbf{V}_2^{\mathcal{T}}] \stackrel{a.s.}{=} \text{rank}[\mathbf{V}_1^{\mathcal{T}}] + \text{rank}[\mathbf{V}_2^{\mathcal{T}}] \\
&\stackrel{\text{(Lemma 5)}}{\leq} \mathbf{r}_1 + \mathbf{r}_2 \\
&\stackrel{\substack{\text{(Lemma 5)}\\ a.s.}}{\leq} \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{V}_2^n] + \text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{V}_1^n] \\
&\stackrel{a.s.}{=} 2\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n] - \text{rank}[\mathbf{G}_{12}^n\mathbf{V}_2^n] \\
&\stackrel{a.s.}{\leq} 2\text{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] - \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n].
\end{aligned}
$$

By rearranging the two sides of the above inequality, the proof of Lemma 1 would be complete.

In the next section we present another application of Rank Ratio Inequality in analyzing fundamental limits of interference management in interference networks. More specifically, by using Rank Ratio Inequality, we derive a new upper bound on linear DoF of three-user interference channel with delayed CSIT.

2.3 The Three-User Interference Channel with Delayed CSIT

In this section we demonstrate how Rank Ratio Inequality can be applied to the three-user interference channel with delayed CSIT depicted in Fig. 2.2. In particular, we derive a new upper bound of $\frac{9}{7}$ on linear DoF of three-user interference channel with delayed CSIT. This is the first upper bound that captures the impact of delayed CSIT on the degrees of freedom of this network. We use the same notation as in the previous section.

The channel model is similar to that of the X-channel except the channel input-output relation and decodability constraints. The received signal at Rx_k

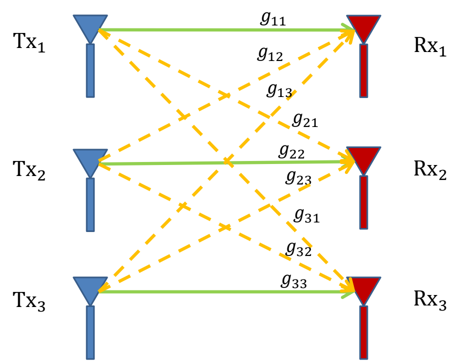


Figure 2.2: Network configuration for the three-user IC. There are three transmitters and three receivers, and for $j = 1, 2, 3$, Tx_j has message for Rx_j. We assume time-varying channels, with delayed CSIT.

$(k \in \{1, 2, 3\})$ at time t is given by

$$\mathbf{y}_k(t) = \sum_{j=1}^{3} \mathbf{g}_{kj}(t)\mathbf{x}_j(t) + \mathbf{z}_k(t). \tag{2.27}$$

For block length of n and $j = 1, 2, 3$, we consider the decodability constraint of

$$\dim\left(\mathrm{Proj}_{\mathcal{I}_j^c}\mathrm{colspan}(\mathbf{G}_{jj}^n\mathbf{V}_j^n)\right) = \dim\left(\mathrm{colspan}(\mathbf{V}_j^n)\right) = m_j(n), \tag{2.28}$$

where $\mathcal{I}_j = \cup_{i \neq j}\mathrm{colspan}(\mathbf{G}_{ji}^n\mathbf{V}_i^n)$. Denote the linear degrees of freedom region $\mathcal{D}_{\text{3UserIC}}$ as the closure of the set of all achievable 3-tuples (d_1, d_2, d_3), where $d_j = \lim_{n\to\infty}\frac{m_j(n)}{n}$, and $\{m_1(n), m_2(n), m_3(n)\}$ are linearly achievable with probability 1 for every $n \in \mathbb{N}$. We are interested in characterizing the sum linear degrees of freedom:

$$\mathrm{DoF}_{\text{L-sum}} = \max \sum_{j=1}^{3} d_j, \qquad s.t. \quad (d_1, d_2, d_3) \in \mathcal{D}. \tag{2.29}$$

With delayed CSIT, it was shown in [66] that the sum DoF of $\frac{9}{8}$ can be achieved, which was later improved to $\frac{36}{31}$ in [6]. However, the best known outer bound so far is $\frac{3}{2}$, which also holds for the case of instantaneous CSIT [16]. The following theorem provides a tighter bound on the linear degrees of freedom.

Theorem 2. *For the three-user interference channel with delayed CSIT,*

$$\mathrm{DoF}_{\text{L-sum}} \leq \frac{9}{7}. \tag{2.30}$$

Proof. Let us denote the symmetric degrees of freedom for three-user interference channel by $\mathrm{DoF}_{\text{L-sym}}$. Note that due to symmetry of topology,

$$\mathrm{DoF}_{\text{L-sum}} = 3 \times \mathrm{DoF}_{\text{L-sym}}. \tag{2.31}$$

Hence, in order to prove the theorem it suffices to show that $\mathrm{DoF}_{\text{L-sym}} \leq \frac{3}{7}$. So assume that for a given block length n, $m_1(n) = m_2(n) = m_3(n)$, and we seek to show that if decodability is accomplished with probability 1, we should have $m_1(n) \leq \frac{3}{7}n$. By Lemma 3 if the decodability constraints in (2.28) are satisfied with probability 1 for pairs Tx_1-Rx_1 and Tx_2-Rx_2, then

$$\mathrm{rank}[\mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n] + \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n] \overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n], \tag{2.32}$$

$$\mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] + \mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n] \overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n], \tag{2.33}$$

where $\mathrm{rank}[\mathbf{V}_1^n] \overset{a.s.}{=} \mathrm{rank}[\mathbf{V}_2^n] \overset{a.s.}{=} \mathrm{rank}[\mathbf{V}_3^n] \overset{a.s.}{=} m_1(n)$. Thus, assuming $m_1(n) = m_2(n) = m_3(n)$ are linearly achievable with probability 1, from (2.33), we have

$$\begin{aligned}\mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n] &\overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] - \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] \\ &\overset{(a)}{\leq} \mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] - \mathrm{rank}[\mathbf{G}_{23}^n\mathbf{V}_3^n],\end{aligned} \tag{2.34}$$

where (a) follows from sub-modularity of rank (Lemma 2). In addition, we know that

$$\mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n] \geq \mathrm{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] - \mathrm{rank}[\mathbf{G}_{23}^n\mathbf{V}_3^n]. \tag{2.35}$$

By (2.34), (2.35) we conclude that

$$\begin{aligned}\text{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] &\overset{a.s.}{=} \text{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n] + \text{rank}[\mathbf{G}_{23}^n\mathbf{V}_3^n] \\ &\overset{a.s.}{=} \text{rank}[\mathbf{V}_2^n] + \text{rank}[\mathbf{V}_3^n] \overset{a.s.}{=} 2m_1(n).\end{aligned} \tag{2.36}$$

On the other hand, from Lemma 6 we know that

$$\text{rank}[\mathbf{G}_{22}^n\mathbf{V}_2^n \quad \mathbf{G}_{23}^n\mathbf{V}_3^n] \overset{a.s.}{\leq} \frac{3}{2}\text{rank}[\mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n]. \tag{2.37}$$

Hence, by (2.36), (2.37),

$$\text{rank}[\mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n] \overset{a.s.}{\geq} \frac{4}{3}m_1(n). \tag{2.38}$$

Finally, by considering (2.32), (2.38), and the fact that $\text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n] \leq n$, we get

$$\begin{aligned}m_1(n) &\overset{a.s.}{=} \text{rank}[\mathbf{V}_1] \overset{a.s.}{=} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n] \\ &\overset{\substack{(2.32)\\a.s}}{=} \text{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n] - \text{rank}[\mathbf{G}_{12}^n\mathbf{V}_2^n \quad \mathbf{G}_{13}^n\mathbf{V}_3^n] \\ &\overset{\substack{(2.38)\\a.s.}}{\leq} n - \frac{4}{3}m_1(n),\end{aligned}$$

which implies that $m_1(n) \leq \frac{3}{7}n$ because $n, m_1(n)$ are non-random, and this completes the proof. □

2.4 Interference Channel with Limited Cooperation and Delayed CSIT

In this section we study the two-user Gaussian interference channel with fading channels and delayed CSIT (Fig. 2.3). We first present the network model and propose a model for capturing transmitter cooperation via quantifying the amount of message sharing between the transmitters. We then completely characterize the DoF of two-user interference channel with delayed CSIT and partial transmitter cooperation, and present the proof.

2.4.1 System Model and Main Results

We consider the Gaussian interference channel depicted in Fig. 2.3. It consists of two transmitters and two receivers, where Tx_i has a message W_i for Rx_i, $i \in \{1, 2\}$. Each node is equipped with a single antenna.

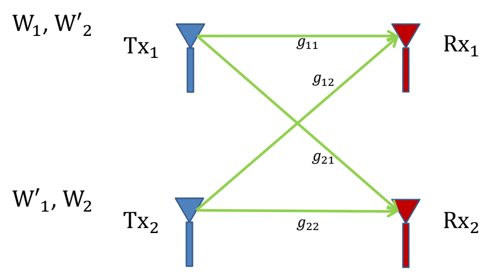


Figure 2.3: Network configuration for interference channel. There are two transmitters and two receivers, where Tx_i has a message W_i for Rx_i, $i \in \{1, 2\}$. In addition, when transmitters cooperate, Tx_1 has access to W_2', and Tx_2 has access to W_1', where W_1', W_2' can be thought of as portions of the messages W_1, W_2, respectively. Moreover, we assume time-varying channels, with delayed CSIT.

The received signal at Rx_j ($j \in \{1, 2\}$) at time t is given by

$$Y_j(t) = G_{j1}(t)X_1(t) + G_{j2}(t)X_2(t) + Z_j(t), \tag{2.39}$$

where $X_i(t)$ is the transmit signal of Tx_i; $G_{ji}(t) \in \mathbb{C}$ indicates coefficient of the channel from Tx_i to Rx_j; and $Z_j(t) \sim \mathcal{CN}(0, 1)$ is the complex additive white Gaussian noise which is i.i.d. across receivers. The channel coefficients $G_{ji}(t)$'s are i.i.d across time and users, and they are drawn from a continuous distribution. We denote by $\mathcal{G}(t)$ the set of all four channel coefficients at time t. In addition, we denote by $\mathcal{G}^n$ the set of all channel coefficients from time 1 to n, i.e.,

$$\mathcal{G}^n = \{G_{ji}(t) : i, j \in \{1,2\}, t = 1, \ldots, n\}.$$

We denote by $X_i^t, Y_j^t, Z_j^t, G_{ji}^t, \mathcal{G}^t$, the transmit signal vector of Tx_i at times $1, 2, \ldots, t$, the received signal vector of Rx_j at times $1, 2, \ldots, t$, the received vector of noise at Rx_j at times $1, 2, \ldots, t$, the block diagonal channel matrix comprising of channel between Tx_i and Rx_j at times $1, 2, \ldots, t$, and the set of all channel coefficients at times $1, 2, \ldots, t$, respectively.

Denoting the vector of transmit signals for Tx_i in a block of length n by X_i^n, each transmitter Tx_i obeys an average power constraint, $\frac{1}{n}E\{\|X_i^n\|^2\} \leq P$. We assume delayed channel state information at the transmitters (CSIT). In other words, at time t, only the states of the past (i.e. $\mathcal{G}^{t-1}$) are known to Tx_1, Tx_2.

We abstract the limited cooperation between the two transmitters via assuming that Tx_1 also has access to a partial message W_2', which is independent of W_1, but $I(W_2; W_2')$ can be strictly positive. Similarly, Tx_2 has access to a partial message W_1', which is independent of W_2, but $I(W_1; W_1')$ can be strictly positive (see Fig. 2.3).

Definition 4. *A code for a communication block length of n of the cooperative Gaussian interference channel with delayed CSIT consists of:*

- *A sequence of encoders*

$$f_{1,t}^{(n)} : W_1 \times W_2' \times \mathcal{G}^{t-1} \longmapsto C$$
$$f_{2,t}^{(n)} : W_2 \times W_1' \times \mathcal{G}^{t-1} \longmapsto C$$

where message W_i is uniformly distributed over $\{1, 2, \ldots, M_i(n)\}$, $i \in \{1, 2\}$.

- *The corresponding decoders $F_1^{(n)}, F_2^{(n)}$ at the receivers, where*

$$\hat{W}_1 = F_1^{(n)}(Y_1^n, \mathcal{G}^n)$$
$$\hat{W}_2 = F_2^{(n)}(Y_2^n, \mathcal{G}^n)$$

- *The error probability of communication is defined to be*

$$p_e^{(n)} = \Pr(W_1 \neq \hat{W}_1 \text{ or } W_2 \neq \hat{W}_2).$$

We now define the degrees of freedom for the interference channel with delayed CSIT and partial transmitter cooperation.

Definition 5. *For two arbitrary values of $0 \leq \rho_1, \rho_2 \leq 1$, degrees of freedom (DoF) pair $(d_1, d_2)_{\rho_1,\rho_2}$ is achievable if there exists a sequence of encoding and decoding schemes for which*

$$\limsup_{n\to\infty} \frac{I(W_i; W_i')}{H(W_i)} \leq \rho_i, \quad i \in \{1, 2\}, \tag{2.40}$$

with

$$\limsup_{n\to\infty} p_e^{(n)} = 0, \tag{2.41}$$

and

$$\liminf_{P\to\infty} \liminf_{n\to\infty} \frac{\log M_i(n)}{n \log P} \geq d_i, \quad i \in \{1, 2\}. \tag{2.42}$$

Further, we define $\mathcal{D}_{\rho_1,\rho_2}$ to be the closure of the set of all achievable DoF pairs $(d_1, d_2)_{\rho_1,\rho_2}$. In addition, sum DoF, or $\mathrm{DoF}_{\rho_1,\rho_2}$, is defined as

$$\mathrm{DoF}_{\rho_1,\rho_2} \triangleq \sup\{d_1 + d_2 \quad | \quad (d_1, d_2) \in \mathcal{D}_{\rho_1,\rho_2}\}. \tag{2.43}$$

In our model, transmitter cooperation is enabled via message sharing between the transmitters. Hence, the network can be viewed as interference channel with cognition, or in other words, cognitive/cooperative interference channel. In fact, one can think of ρ_1, ρ_2 as "cooperation fractions". Roughly speaking,

ρ_1 is the fraction of message W_1 which is available to Tx_2. Note that for the special case of $\rho_1 = \rho_2 = 0$, $\mathcal{D}_{\rho_1,\rho_2}$ is the DoF region for the two-user interference channel with delayed CSIT. On the other hand, for the case of $\rho_1 = \rho_2 = 1$, $\mathcal{D}_{\rho_1,\rho_2}$ is the DoF region for the two-user multiple-input single-output (MISO) broadcast channel with delayed CSIT (see Fig. 2.4). Therefore, interference channel and MISO broadcast channel can be viewed as two extreme cases of cooperation between transmitters, where in the former there is no information shared between the transmitters about the messages, while in the latter the transmitters both have access to both messages W_1, W_2. So, a fundamental question is: how does the DoF change with the amount of information that is shared between the transmitters? The following Theorem bridges the gap between no cooperation and full cooperation, and characterizes the DoF region as a function of ρ_1, ρ_2.

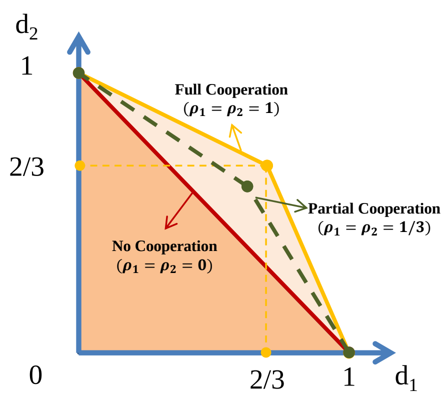


Figure 2.4: DoF regions for three cases of transmitter cooperation. For the special case of $\rho_1 = \rho_2 = 0$, $\mathcal{D}_{\rho_1,\rho_2}$ is the DoF region for the two-user interference channel with delayed CSIT, which is characterized by $d_1 + d_2 \leq 1$. On the other hand, for the special case of $\rho_1 = \rho_2 = 1$, $\mathcal{D}_{\rho_1,\rho_2}$ is the DoF region for the two-user multiple-input single-output (MISO) broadcast channel with delayed CSIT, which is characterized by two inequalities $d_1 + \frac{d_2}{2} \leq 1$ and $\frac{d_1}{2} + d_2 \leq 1$. The third DoF region, which corresponds to $\rho_1 = \rho_2 = \frac{1}{3}$, contains that of no cooperation, but is contained by that of full cooperation.

Theorem 3. *The DoF region for the two-user interference channel with delayed CSIT* *and limited transmitter cooperation is characterized as follows:*

$$\mathcal{D}_{\rho_1,\rho_2} = \left\{ (d_1, d_2) \quad \Bigg| \quad \begin{array}{c} d_1 + \max(\frac{1}{2}, 1-\rho_2) d_2 \leq 1 \\ \max(\frac{1}{2}, 1-\rho_1) d_1 + d_2 \leq 1 \end{array} \right\}. \tag{2.44}$$

A direct consequence of the result in Theorem 3 is the following Corollary.

Corollary 1.

$$\mathrm{DoF}_{\rho_1,\rho_2} = \frac{2 - \max(\frac{1}{2}, 1-\rho_2) - \max(\frac{1}{2}, 1-\rho_1)}{1 - \max(\frac{1}{2}, 1-\rho_2) \times \max(\frac{1}{2}, 1-\rho_1)}, \tag{2.45}$$

which for $\rho_1 = \rho_2 = \rho$ results in the following:

$$\mathrm{DoF}_{\rho,\rho} = \frac{2}{\max(\frac{3}{2}, 2-\rho)}. \tag{2.46}$$

The above result is illustrated in Fig. 2.5.

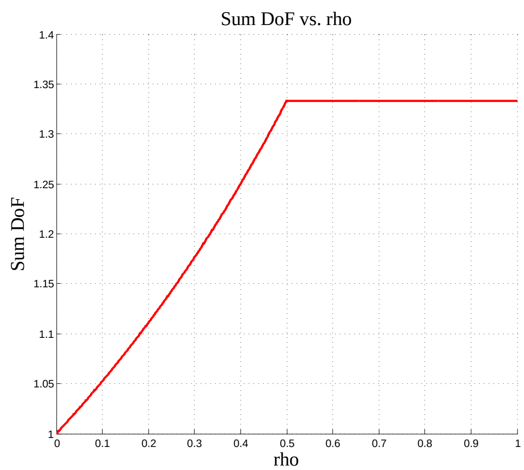


Figure 2.5: Sum DoF for the two-user interference channel with delayed CSIT and transmitter cooperation as a function of the cooperation fraction ρ, where $\rho = \rho_1 = \rho_2$.

Proposition 1. *Corollary 1 implies that $\rho_1 = \rho_2 = \frac{1}{2}$ is sufficient to essentially turn the network into the two-user MISO broadcast channel with delayed CSIT. Therefore, the two transmitters in an interference channel with delayed CSIT do not need to share their entire messages with one another in order to benefit from the DoF gains; as long as half of the messages are shared, the network can enjoy the capabilities of a MISO broadcast channel.*

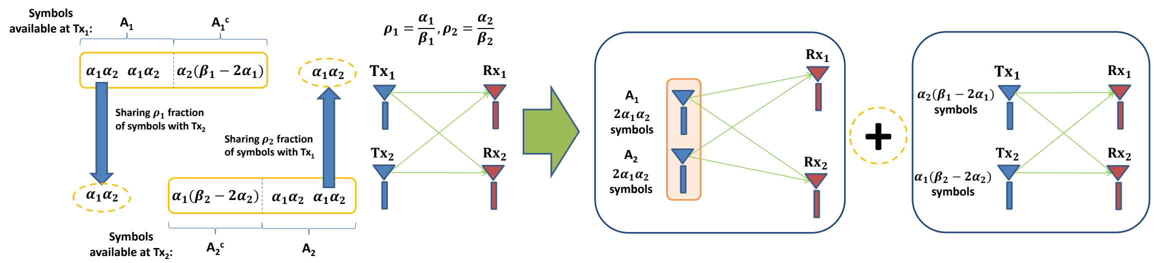


Figure 2.6: Interference channel with delayed CSIT and limited transmitter cooperation $0 \leq \rho_1, \rho_2 \leq \frac{1}{2}$, where $\rho_1 = \frac{\alpha_1}{\beta_1}, \rho_2 = \frac{\alpha_2}{\beta_2}$. Assuming Tx_1 has $\alpha_2\beta_1$ symbols to deliver to Rx_1, and Tx_2 has $\alpha_1\beta_2$ symbols to deliver to Rx_2, the network can be viewed as if it is comprised of a MISO broadcast channel, together with an interference channel (with a different weight for each). Similarly, the achievable scheme consists of a phase in which the network essentially operates as a broadcast channel, and a phase in which TDMA is used for delivering symbols (i.e. the network operates as an interference channel with no cooperation). See Section 2.4.2 for details of the achievable scheme.

2.4.2 Proof of Achievability for Theorem 3

In this section we first make the following observation, which will be an important ingredient of the achievable schemes.

Proposition 2. *Consider a two-user interference channel with delayed CSIT, where* Tx_i *($i \in \{1,2\}$) has $2m$ information symbols to deliver to* Rx_i *($m \in \mathbb{N}$); and* Tx_i *also has access to half of the symbols available to* Tx_{3-i} *which are to be delivered to* Rx_{3-i}*. Then, all the $4m$ information symbols can be delivered to their corresponding receivers in $3m$ timeslots.*

To see why Proposition 2 holds, it is sufficient to show its validity for $m = 1$. So we would only need to show that if Tx_1 has 2 symbols to deliver to Rx_1, where one of those symbols is shared with Tx_2, and vice versa for Tx_2, then we can deliver all 4 symbols over 3 timeslots. But this can be done simply by applying MAT scheme [65] developed for two-user MISO broadcast channel with delayed CSIT.

We now prove the achievability for Theorem 3. Note that due to the structure of the polytope proposed by Theorem 3 for the DoF region, there is only one non-trivial extreme point (i.e. one extreme point other than $(0,0),(0,1),(1,0)$). That extreme point, denoted by (d_1^*, d_2^*), is as follows:

$$d_1^* = \frac{1 - \max(\frac{1}{2}, 1-\rho_2)}{1 - \max(\frac{1}{2}, 1-\rho_1) \times \max(\frac{1}{2}, 1-\rho_2)},$$
$$d_2^* = \frac{1 - \max(\frac{1}{2}, 1-\rho_1)}{1 - \max(\frac{1}{2}, 1-\rho_1) \times \max(\frac{1}{2}, 1-\rho_2)}. \tag{2.47}$$

Therefore, to prove the achievability it is sufficient to provide a scheme that achieves (d_1^*, d_2^*). For ease of exposition, we divide the range of values of ρ_1, ρ_2 into 4 categories, and show how (d_1^*, d_2^*) can be achieved in each category.

$1 \geq \rho_1, \rho_2 \geq \frac{1}{2}$

Note that for $\rho_1, \rho_2 \geq \frac{1}{2}$, $d_1^* = d_2^* = \frac{2}{3}$ by (2.47). Also, note that increasing ρ_1, ρ_2 can only improve the DoF region. Therefore, it is sufficient to show the achievability of $(d_1^*, d_2^*) = (\frac{2}{3}, \frac{2}{3})$ for $\rho_1 = \rho_2 = \frac{1}{2}$. But, the network considered in Proposition 2 is already an instance with $\rho_1 = \rho_2 = \frac{1}{2}$, where $(d_1, d_2) = (\frac{2}{3}, \frac{2}{3})$ can be achieved.

$0 \leq \rho_1, \rho_2 \leq \frac{1}{2}$

Let us first consider the case that ρ_1, ρ_2 are rational numbers; i.e. there exist $\alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbb{N}$ such that $\rho_1 = \frac{\alpha_1}{\beta_1}, \rho_2 = \frac{\alpha_2}{\beta_2}$. Note that when $0 \leq \rho_1, \rho_2 \leq \frac{1}{2}$, by (2.47) we have

$$(d_1^*, d_2^*) = \left(\frac{\rho_2}{\rho_1 + \rho_2 - \rho_1\rho_2}, \frac{\rho_1}{\rho_1 + \rho_2 - \rho_1\rho_2}\right)$$

$$= \left(\frac{\alpha_2\beta_1}{\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_1\alpha_2}, \frac{\alpha_1\beta_2}{\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_1\alpha_2}\right). \tag{2.48}$$

Let us provide $\alpha_2\beta_1$ symbols to Tx_1 to be delivered to Rx_1, where $\alpha_1\alpha_2$ of those symbols are shared with Tx_2. In addition, we provide $\alpha_1\beta_2$ symbols to Tx_2 to be delivered to Rx_2, where $\alpha_1\alpha_2$ of those symbols are shared with Tx_1. In such scenario, the fractions of message sharing comply with the ratios ρ_1, ρ_2; since $\frac{\alpha_1\alpha_2}{\alpha_2\beta_1} = \frac{\alpha_1}{\beta_1} = \rho_1$, and $\frac{\alpha_1\alpha_2}{\alpha_1\beta_2} = \frac{\alpha_2}{\beta_2} = \rho_2$. So, by (2.48) it is sufficient to deliver all the symbols over $(\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_1\alpha_2)$ timeslots.

We choose $2\alpha_1\alpha_2$ symbols intended for Rx_1 that are available to Tx_1[5], which include the $\alpha_1\alpha_2$ symbols shared with Tx_2. Denote the set of these $2\alpha_1\alpha_2$ symbols by A_1; and denote the remaining $\alpha_2\beta_1 - 2\alpha_2\alpha_1$ symbols available to Tx_1 by A_1^c.

We also choose $2\alpha_1\alpha_2$ symbols intended for Rx_2 that are available to Tx_2, which include the $\alpha_1\alpha_2$ symbols shared with Tx_1. Denote the set of these $2\alpha_1\alpha_2$ symbols by A_2; and denote the remaining $\alpha_1\beta_2 - 2\alpha_2\alpha_1$ symbols available to Tx_2 by A_2^c.

Using Proposition 2 one can deliver the symbols in A_1, A_2 in $3\alpha_1\alpha_2$ timeslots. Furthermore, the $\alpha_2\beta_1 - 2\alpha_2\alpha_1$ symbols in A_1^c and $\alpha_1\beta_2 - 2\alpha_2\alpha_1$ symbols in A_2^c can be simply delivered by TDMA over $(\alpha_2\beta_1 - 2\alpha_2\alpha_1) + (\alpha_1\beta_2 - 2\alpha_2\alpha_1)$ timeslots. Hence, over a total of $\alpha_2\beta_1 + \alpha_1\beta_2 - \alpha_2\alpha_1$ timeslots all the symbols can be delivered. (Refer to Fig. 2.6 for an illustration of the achievable scheme.)

For the case that ρ_1, ρ_2 are not necessarily rational numbers, since rationals are dense in real numbers, one can consider an increasing sequence $\{\rho_1(n), \rho_2(n)\}_{n=1}^{\infty}$, where $\lim_{n\to\infty} \rho_i(n) = \rho_i$, $i = 1, 2$; and then apply the above achievable scheme for each n to asymptotically achieve (d_1^*, d_2^*).

[5]Note that $\alpha_2\beta_1 \geq 2\alpha_1\alpha_2$; therefore, we can always choose such a subset of symbols.

$0 \le \rho_1 \le \frac{1}{2}, \rho_2 \ge \frac{1}{2}$

Note that when $\rho_1 = \frac{\alpha_1}{\beta_1}, \rho_2 = \frac{\alpha_2}{\beta_2}$ and $0 \le \rho_1 \le \frac{1}{2}, \rho_2 \ge \frac{1}{2}$, by (2.47) we have

$$(d_1^*, d_2^*) = (\frac{1}{1+\rho_1}, \frac{2\rho_1}{1+\rho_1}) = (\frac{\beta_1}{\alpha_1+\beta_1}, \frac{2\alpha_1}{\alpha_1+\beta_1}). \tag{2.49}$$

We provide β_1 symbols to Tx_1 to be delivered to Rx_1, where α_1 of those symbols are shared with Tx_2. In addition, we provide $2\alpha_1$ symbols to Tx_2 to be delivered to Rx_2, where α_1 of those symbols are shared with Tx_1. In such scenario, the message sharing ratios comply with the values of ρ_1, ρ_2. So, by (2.49) it is sufficient to deliver all the symbols over $\alpha_1 + \beta_1$ timeslots.

Similar to the case of $0 \le \rho_1, \rho_2 \le \frac{1}{2}$, we group the symbols into subsets A_1, A_1^c, A_2, A_2^c, where in this case, $|A_1| = 2\alpha_1, |A_1^c| = \beta_1 - 2\alpha_1, |A_2| = 2\alpha_1, |A_2^c| = 0$. Therefore, we spend $3\alpha_1$ to deliver the symbols in A_1, A_2, and spend an additional $\beta_1 - 2\alpha_1$ timeslots to deliver symbols in A_1^c. Hence, all symbols can be delivered over $\alpha_1 + \beta_1$ timeslots.

$\rho_1 \ge \frac{1}{2}, 0 \le \rho_2 \le \frac{1}{2}$

This case is very similar to the case of $0 \le \rho_1 \le \frac{1}{2}, \rho_2 \ge \frac{1}{2}$; therefore, the proof is omitted for brevity.

2.4.3 Proof of Converse for Theorem 3

In order to prove the converse, we first state the following Lemma (Lemma 6) which is the key ingredient of the converse. We then show how Lemma 6 is used to complete the proof of Theorem 3; and finally, we prove Lemma 6.

Lemma 6. *Consider the two-user interference channel with transmitter cooperation as defined in Section 5.2, which satisfies (2.40)-(2.42) for $i = 1$. Then,*

$$h(Y_2^n|W_2, \mathcal{G}^n) + n \times o(\log P) \geq (1 - \rho_1)h(Y_1^n|W_2, \mathcal{G}^n). \tag{2.50}$$

Remark 3. *Lemma 6 indicates that once d_1 DoF is delivered to* Rx_1*, then at least $(1 - \rho_1)d_1$ dimensions are occupied at* Rx_2 *by the signal intended for* Rx_1 *(which is interference to* Rx_2*). One can view Lemma 6 as an "entropy ratio inequality", similar to the Rank Ratio Inequalities developed in [52, 51, 47, 46, 48]. However, there are differences between Lemma 6 and Rank Ratio Inequality in [52]; in the Rank Ratio Inequality, the objective is to bound the maximum ratio of dimensions of received signals at the two receivers, while transmitters cannot cooperate, and can only use linear encoding schemes. In addition, in the Rank Ratio Inequality there is no decodability assumption. Nevertheless, Lemma 6 assumes the decodability of W_1 at Rx_1[6], considers transmitter cooperation, and is not restricted to linear encoding schemes.*

We now use Lemma 6 to prove the converse. Intuitively, since at least $(1 - \rho_1)d_1$ dimension is occupied at Rx_2 by the signal intended for Rx_1, by dimension counting at Rx_2 we arrive at $(1 - \rho_1)d_1 + d_2 \leq 1$.

Proof. Consider arbitrary but fixed ρ_1, ρ_2, where $0 \leq \rho_1, \rho_2 \leq 1$. Note that the outer bounds provided in [65] for 2-user MISO broadcast channel with delayed CSIT are also valid for the 2-user interference channel with transmitter cooperation and delayed CSIT. Therefore, we have

$$d_1 + \frac{1}{2}d_2 \leq 1, \tag{2.51}$$

$$\frac{1}{2}d_1 + d_2 \leq 1. \tag{2.52}$$

[6]The same inequality would not necessarily hold once the decodability assumption is removed.

Hence, to prove the converse for Theorem 3 it is sufficient to prove that $d_1 + (1 - \rho_2)d_2 \leq 1$, and $(1-\rho_1)d_1 + d_2 \leq 1$. We only prove the latter here; as the former can be proven similarly, and hence its proof is omitted for brevity.

Suppose the DoF pair (d_1, d_2) is achievable. Then, there exists a sequence of coding schemes for which the conditions (2.40)-(2.42) are satisfied. By Fano's inequality we have

$$
\begin{aligned}
n(R_2 - o(\log P)) \leq & I(W_2; Y_2^n|\mathcal{G}^n) = h(Y_2^n|\mathcal{G}^n) - h(Y_2^n|W_2, \mathcal{G}^n) \\
\overset{\text{(Lemma 6)}}{\leq} & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)h(Y_1^n|W_2, \mathcal{G}^n) + n \times o(\log P) \\
= & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)[h(Y_1^n|W_2, \mathcal{G}^n) - h(Y_1^n|W_1, W_2, \mathcal{G}^n)] \\
& + (1-\rho_1)h(Y_1^n|W_1, W_2, \mathcal{G}^n) + n \times o(\log P) \\
\overset{(a)}{=} & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)I(W_1; Y_1^n|W_2, \mathcal{G}^n) \\
& + (1-\rho_1)h(Z_1^n|W_1, W_2, \mathcal{G}^n) + n \times o(\log P) \\
\overset{(b)}{=} & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)I(W_1; Y_1^n|W_2, \mathcal{G}^n) + n \times o(\log P) \\
= & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)H(W_1|W_2, \mathcal{G}^n) \\
& + (1-\rho_1)H(W_1|W_2, Y_1^n, \mathcal{G}^n) + n \times o(\log P) \\
\overset{\text{(Fano's)}}{\leq} & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)H(W_1|W_2, \mathcal{G}^n) + n \times o(\log P) \\
\overset{(c)}{=} & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)H(W_1) + n \times o(\log P) \\
= & h(Y_2^n|\mathcal{G}^n) - (1-\rho_1)nR_1 + n \times o(\log P) \\
= & \sum_{i=1}^{n} h(Y_2(i)|Y_2^{i-1}, \mathcal{G}^n) - (1-\rho_1)nR_1 + n \times o(\log P) \\
\leq & n \times \log P - (1-\rho_1)nR_1 + n \times o(\log P),
\end{aligned}
\tag{2.53}
$$

where (a) holds since $W_1, W_2, \mathcal{G}^n$ determine $G_{11}^n X_1^n + G_{12}^n X_2^n$; (b) holds since $h(Z_1^n|W_1, W_2, \mathcal{G}^n) \overset{\text{(independence)}}{=} h(Z_1^n) = n \times \log 2\pi e = n \times o(\log P)$; and (c) holds since W_1 is independent of $W_2, \mathcal{G}^n$.

Therefore, by rearranging and dividing both sides of (2.53) by $n \times \log P$, and taking the limit of both $P, n \to \infty$, we obtain

$$(1-\rho_1)d_1 + d_2 \leq 1, \tag{2.54}$$

which completes the converse proof. $\square$

We now prove Lemma 6. We have,

$$\begin{aligned}
(1-\rho_1)h(Y_1^n|W_2,\mathcal{G}^n) &= (1-\rho_1)I(W_1;Y_1^n|W_2,\mathcal{G}^n) + (1-\rho_1)h(Y_1^n|W_1,W_2,\mathcal{G}^n) \\
&= (1-\rho_1)I(W_1;Y_1^n|W_2,\mathcal{G}^n) + (1-\rho_1)h(Z_1^n|W_1,W_2,\mathcal{G}^n) \\
&= (1-\rho_1)I(W_1;Y_1^n|W_2,\mathcal{G}^n) + n \times o(\log P) \\
&\leq (1-\rho_1)H(W_1|W_2,\mathcal{G}^n) + n \times o(\log P) \\
&\overset{(d)}{=} H(W_1) - \rho_1 H(W_1) + n \times o(\log P) \\
&\overset{(2.40)}{\leq} H(W_1) - I(W_1;W_1') + n \times o(\log P) \\
&\overset{(e)}{=} H(W_1|W_1',W_2,\mathcal{G}^n) + n \times o(\log P) \\
&= I(W_1;Y_2^n|W_1',W_2,\mathcal{G}^n) + H(W_1|Y_2^n,W_1',W_2,\mathcal{G}^n) + n \times o(\log P) \\
&\leq h(Y_2^n|W_1',W_2,\mathcal{G}^n) + H(W_1|Y_2^n,W_1',W_2,\mathcal{G}^n) + n \times o(\log P) \\
&\leq h(Y_2^n|W_2,\mathcal{G}^n) + H(W_1|Y_2^n,W_1',W_2,\mathcal{G}^n) + n \times o(\log P) \qquad (2.55) \\
&\overset{(f)}{=} h(Y_2^n|W_2,\mathcal{G}^n) + H(W_1|X_2^n,G_{21}^nX_1^n+Z_2^n,Y_2^n,W_1',W_2,\mathcal{G}^n) + n \times o(\log P) \\
&\overset{(g)}{=} h(Y_2^n|W_2,\mathcal{G}^n) + H(W_1|X_2^n,G_{11}^nX_1^n+Z_1^n,\mathcal{G}^n) + n \times o(\log P) \\
&= h(Y_2^n|W_2,\mathcal{G}^n) + H(W_1|Y_1^n,X_2^n,G_{11}^nX_1^n+Z_1^n,\mathcal{G}^n) + n \times o(\log P) \\
&\leq h(Y_2^n|W_2,\mathcal{G}^n) + H(W_1|Y_1^n,\mathcal{G}^n) + n \times o(\log P) \\
&\overset{(\text{Fano's})}{\leq} h(Y_2^n|W_2,\mathcal{G}^n) + n \times o(\log P), \qquad (2.56)
\end{aligned}$$

where (d) holds since W_1 is independent of $(W_2,\mathcal{G}^n)$; (e) holds due to the Markov chain $W_1 \leftrightarrow W_1' \leftrightarrow (W_2,\mathcal{G}^n)$; (f) holds since $W_1', W_2, \mathcal{G}^n$ determine X_2^n; therefore, by

$X_2^n, \mathcal{G}^n, Y_2^n$ one can determine $G_{21}^n X_1^n + Z_2^n$ by subtracting $G_{22}^n X_2^n$ from Y_2^n; and (g) holds due to noise and channel symmetry.

Remark 4. *An important observation in the proof of Lemma 6 is that removing W_1' from the conditioning in (2.55) does not make the upper bound loose. The intuition originates from the achievability, in which for a DoF optimal scheme the signals corresponding to the message W_1 will occupy a dimension at* Rx_2 *which roughly equals $H(W_1|W_1')$, and not $H(W_1)$. This is due to the interference alignment occurring at* Rx_2*, which is in turn due to W_1' being shared with* Tx_2.

Remark 5. *Proof of Lemma 6 does not rely on the type of CSIT available to the transmitters. Therefore, one can use Lemma 6 for any type of CSIT. As a result, Lemma 6 can be used to analyze the DoF for two-user interference channel with instantaneous CSIT and limited transmitter cooperation.*

Proposition 3. *The sum DoF for the two-user interference channel with instantaneous CSIT and limited transmitter cooperation is characterized as follows:*

$$\mathrm{DoF}_{\rho_1,\rho_2} = \frac{\rho_1 + \rho_2}{1 - (1 - \rho_2) \times (1 - \rho_1)}, \tag{2.57}$$

which for $\rho_1 = \rho_2 = \rho$ results in the following:

$$\mathrm{DoF}_{\rho,\rho} = \frac{2}{2 - \rho}. \tag{2.58}$$

Proof of the above Proposition uses the same techniques used in the achievability and converse proof for the case of delayed CSIT; therefore, it is omitted for brevity.

2.5 $2 \times k$ MISO Broadcast Channel with Delayed CSIT

In this section we study the problem of k-user MISO Broadcast Channel, where there are k single-antenna receivers and a 2-antenna transmitter, as depicted in Fig. 2.7. The transmitter has access to the delayed CSIT; hence, the problem is called the $2 \times k$ MISO BC with delayed CSIT.

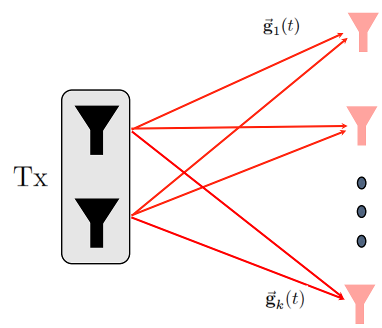


Figure 2.7: Network configuration for $2 \times k$-user MISO BC with delayed CSIT.

For this problem there have been lower and upper bounds developed on its DoF in [65]. However, the bounds are only tight for the special cases of $k = 2, 3$, and beyond that, there is a clear gap between the existing lower and upper bounds.

In this work, we first provide a new achievable scheme for 2×4 MISO BC with delayed CSIT, which improves the state-of-the-art in [65]. We then provide a generalization of our achievable scheme for $2 \times k$ MISO BC with delayed CSIT.

2.5.1 New Achievable Scheme for 2×4 MISO BC with Delayed CSIT

We provide a new achievable scheme that achieves $\frac{14}{9}$ for 2×4 MISO BC with delayed CSIT. The scheme consists of 4 phases.

Phase 1:

The duration of Phase 1 is 6 time slots. In each time slot 4 new symbols are transmitted. Let us denote the symbols desired by $\mathrm{Rx}_1, \mathrm{Rx}_2, \mathrm{Rx}_3, \mathrm{Rx}_4$ by a_i, b_i, c_i, d_i, where $i \in \mathbb{N}$. In particular, in each time slot 2 of the 4 receivers are chosen, and for each, 2 symbols are transmitted. For instance, for the first time slot $\mathrm{Rx}_1, \mathrm{Rx}_2$ are chosen, and the first transmit antenna transmits $a_1 + b_1$ while the second transmit antenna transmits $a_2 + b_2$. As a result, the 4 receivers receive linear equations about a_1, b_1, a_2, b_2, which can be denoted as $A_j + B_j$ for Rx_j. Note that A_2, B_1 both are desired by both $\mathrm{Rx}_1, \mathrm{Rx}_2$. Therefore, A_2, B_1 can be viewed as order-2 symbols. A similar transmission occurs in the next 5 time slots; and similar to time slot 1, at the end of each time slot 2 order-2 symbols are created. Therefore, there will be a total of 12 order-2 symbols created by the end of Phase 1. Hence, if we define the DoF for delivering order-j symbols in the 2×4 MISO BC with delayed CSIT by $DoF_j(2,4)$, we have

$$DoF_1(2,4) \geq \frac{24}{6 + \frac{12}{DoF_2(2,4)}}. \tag{2.59}$$

Phase 2:

In Phase 2 we use 12 time slots to transmit 48 order-2 symbols. In particular, in each time slot we choose 2 types of order-2 symbols that have a desired receiver in common. For instance, in the first time slot we choose order-2 types ab, ac, where the former is desired by Rx_1, Rx_2, and the latter type is desired by Rx_1, Rx_3, and therefore they have Rx_1 in common. More specifically, in the first time slot of Phase 2, first transmit antenna transmits $ab_1 + ac_1$ and the second antenna transmits $ab_2 + ac_2$. As a result, for each $j = 1, 2, 3, 4$, Rx_j receives an equation of the form $AB_j + AC_j$. Note that AB_3 is desired by $\text{Rx}_1, \text{Rx}_2, \text{Rx}_3$, hence can be viewed as an order-3 symbol. Similarly, AC_2 is desired by $\text{Rx}_1, \text{Rx}_2, \text{Rx}_3$, hence can be viewed as an order-3 symbol. In addition, $AB_4 + AC_4$ is known at Rx_4 and desired by Rx_1. Therefore, it can be added to an equation which is desired by Rx_4 and known by Rx_1 to form an order-2 symbol. Hence, $AB_4 + AC_4$ can be viewed as $\frac{1}{2}$ of an order-2 symbol.

Overall, Phase 2 takes 12 time slots, and as a result creates 24 order-3 symbols, and 6 new order-2 symbols. Hence, we have

$$DoF_2(2,4) \geq \frac{48}{12 + \frac{24}{DoF_3(2,4)} + \frac{6}{DoF_2(2,4)}}. \tag{2.60}$$

Phase 3:

For delivering order-3 symbols one can use [65] to achieve $\frac{8}{7}$. In particular, we use 4 time slots, and in each time slot we send one order-3 symbol on each transmit antenna. For instance, in the first time slot of Phase 3 we transmit abc_1

and abc_2. As a result, ABC_4 would be of interest to the other 3 receivers. Hence, by the end of the 4 time slots we can form three linearly independent equations of $ABC_4, ABD_3, ACD_2, BCD_1$ which are of interest to all 4 receivers, and therefore can be viewed as order-4 symbols.

$$DoF_3(2,4) \geq \frac{8}{4 + \frac{3}{DoF_4(2,4)}}. \tag{2.61}$$

Phase 4:

In this Phase we simply send one order-4 symbol during each time slot. Hence,

$$DoF_4(2,4) \geq 1. \tag{2.62}$$

By putting inequalities (2.59)-(2.62) we obtain

$$DoF_1(2,4) \geq \frac{14}{9}, \tag{2.63}$$

$$DoF_2(2,4) \geq \frac{14}{11}, \tag{2.64}$$

which proves achievability of $\frac{14}{9}$ for DoF of 2×4 MISO BC with delayed CSIT. This value is strictly larder than $\frac{3}{2}$ achieved for 2×3 MISO BC with delayed CSIT. However, it is smaller than the upper bound derived using genie-aided arguments and channel enhancement which lead to $\frac{8}{5}$.

In the next Section, we extend the achievable scheme to the general $2 \times k$ MISO BC with delayed CSIT.

2.5.2 Generalization of Our Achievable Scheme to $2 \times k$ MISO BC with delayed CSIT

For the case of general k, we provide a recursive scheme. We first describe the scheme for arbitrary Phase j where $j < k - 1$.

Phase j where $j < k - 1$:

For each time slot of this phase, we send 4 order-j symbols, 2 on each antenna. In particular, we choose 2 types of order-j symbols which have $j - 1$ desired receivers in common. As an example, the first antenna transmits $u_1 \dots u_j + u_1 \dots u_{j-1} u_{j+1}$. As a result of such transmission, each receiver Rx_i receives $(U_1 \dots U_j)_i + (U_1 \dots U_{j-1} U_{j+1})_i$. Therefore, both $(U_1 \dots U_j)_{j+1}$ and $(U_1 \dots U_{j-1} U_{j+1})_j$ are desired by all the first $j + 1$ receivers, which means they can both be viewed as order-$(j + 1)$ symbols. Furthermore, the received signal at Rx_{j+2} is desired by all the first $j - 1$ receivers. Therefore, we can consider it together with $j - 1$ other such equations, and create $j - 1$ equations out of them where each is desired by j receivers. Hence, we have

$$DoF_j(2, k) \geq \frac{4}{1 + \frac{2}{DoF_{j+1}(2,k)} + \frac{\frac{j-1}{j}}{DoF_j(2,k)}}. \tag{2.65}$$

Phase $k - 1$:

For this phase we send two order-$(k - 1)$ symbols of the same type in each time slot, for a total of k time slots. As a result of such transmissions, we can create

$k-1$ order-k symbols, which can in turn be delivered over $k-1$ time slots. Hence,

$$DoF_{k-1}(2,k) \geq \frac{2k}{k+(k-1)}. \tag{2.66}$$

2.6 Concluding Remarks and Future Directions

In this chapter we studied the fundamental limits of interference management in networks where only delayed CSIT is available. In particular, we developed new tools that help better understand and analyze the dynamics of networks when receivers only supply delayed CSIT. First, we focused on X-channel, which is one of the canonical interference networks, and studied the impact of delayed CSIT on its degrees of freedom. We characterized the linear degrees of freedom of the X-channel with delayed CSIT by developing a general lemma (i.e. Rank Ratio Inequality) that shows that, if two distributed transmitters employ linear strategies, the ratio of the dimensions of received linear subspaces at the two receivers cannot exceed $\frac{3}{2}$, due to lack of instantaneous knowledge of the channels. We also applied Rank Ratio Inequality to the three-user interference channel with delayed CSIT, thereby deriving a new upper bound of $\frac{9}{7}$ on its linear degrees of freedom. It is worth mentioning that the result on X-channel with delayed CSIT in this chapter has recently been extended [38] to the MIMO setting by using the techniques developed in this chapter.

We also considered the two-user interference channel with delayed CSIT, and we aimed at answering the following question: how much gain does transmitter cooperation provide in the two-user interference channel with delayed CSIT? To this aim, we first presented a model for capturing the cooperation between the transmitters in terms of the amount of information given to each transmitter about the message available to the other transmitter. We then characterized how the degrees of freedom region of two-user interference channel with delayed CSIT changes as a function of the amount of cooperation between the transmitters.

Finally, we studied interference management with delayed CSIT in the context of 2-by-k MISO broadcast channel with delayed CSIT. We first provided a new achievable scheme for 2-by-4 MISO BC with delayed CSIT, which strictly improves the state-of-the-art achievable scheme. We then generalized the achievable scheme to the k-receiver setting (i.e., 2-by-k MISO BC).

There are various future directions that one can consider with regard to the work done in this chapter. For instance, we conjecture the following generalization of Rank Ratio Inequality (Lemma 1) for general encoding strategies.

Conjecture 1. *Consider the 2-transmitter 2-receiver network setting of Lemma 1. For any $n \in \mathbb{N}$ and any coding strategy denoted by encoding functions $\{f_1^{(n)}, f_2^{(n)}\}$, and its corresponding received signals, $\vec{\mathbf{y}}_1^n$ and $\vec{\mathbf{y}}_2^n$, we have*

$$h(\vec{\mathbf{y}}_1^n|\mathcal{G}^n) \leq \frac{3}{2}h(\vec{\mathbf{y}}_2^n|\mathcal{G}^n) + n \times o(\log(P)). \tag{2.67}$$

Therefore, a future direction would be to remove the linearity restriction on the encoding schemes, and prove (or disprove) the above conjecture, which (if true) will lead to the DoF characterization of the X-channel with delayed CSIT, and a new DoF upper bound for three-user interference channel with delayed CSIT.

We also believe that similar techniques could be applied to other important network configurations to gain insight on how delayed CSIT can be used to improve the Degrees of Freedom, and what the limitations on this DoF improvement are. In particular the k-user interference channel and multi-hop interference networks (e.g., [78, 5, 4]), in which there is a large gap between the state-of-the-art inner and outer bounds on DoF with delayed CSIT, can be considered.

For the problem of interference channel with limited transmitter cooperation and delayed CSIT, a future direction would be to extend the results to the MIMO case, and characterize how transmitter cooperation can improve DoF for MIMO interference channel with delayed CSIT. Another direction is to consider receiver cooperation as well in the two-user interference channel with delayed CSIT.

Finally, for the problem of 2-by-k MISO broadcast channel with delayed CSIT, the DoF is still unsolved for $k > 3$; thus, an interesting future direction is to study whether it is the achievable scheme or the existing genie-aided upper bounds that need to be improved.

CHAPTER 3

INFORMATION-THEORETIC SECURITY WITH PRACTICAL CSIT CONSTRAINTS

3.1 Overview

Wiretap channel is one of the canonical settings in the information-theoretic study of secrecy in communication networks.[1] It consists of a transmitter that wishes to communicate a secret message to a legitimate receiver in the presence of eavesdropper(s) that should not decode the confidential message.

In this chapter, we consider the Gaussian wiretap channel with time-varying channels, where the transmitter is *blind* with respect to the state of channels to eavesdroppers, and only has access to *delayed* channel state information (CSI) of the legitimate receiver, which is referred to as "blind wiretap channel with delayed CSIT". In this setting, we study two important problems. We first study blind multiple-input multiple-output multiple-eavesdropper (MIMOME) wiretap channel with delayed CSIT, where each node in the network is equipped with arbitrary number of antennas.

We completely characterize the secure degrees of freedom (SDoF) of blind MIMOME wiretap channel with delayed CSIT for all antenna configurations. In particular, we strictly improve the state-of-the-art achievable scheme for this network by proposing more efficient artificial noise alignment at the eavesdroppers. Furthermore, we develop a tight upper bound by utilizing four key inequalities that provide lower bounds on the received signal dimensions at re-

[1]The results presented in this chapter are in part based on [47, 46, 49, 56].

ceivers which supply delayed CSIT or no CSIT, or at a collection of receivers where some supply no CSIT. These inequalities together allow for analysis of signal dimensions in networks with asymmetric CSIT.

We then consider a different setting for blind wiretap channel with delayed CSIT, blind cooperative SISO wiretap channel with delayed CSIT, where the secure communication is aided via a distributed jammer, and all nodes in the network are equipped with single antenna. We completely characterize SDoF when only linear coding strategies are employed at the transmitters. As a result, we show that under linear encoding schemes a strictly positive SDoF of $\frac{1}{3}$ is achievable and is optimal. The converse proof is based on Rank Ratio Inequality (Lemma 1), presented in Chapter 2, and Least Alignment Lemma, which implies that once the transmitters in a network have no CSI with respect to a receiver, the least amount of alignment will occur at that receiver, meaning that transmit signals will occupy the maximal signal dimensions at that receiver.

3.2 Blind MIMOME Wiretap Channel with Delayed CSIT

The most fundamental network configuration in information-theoretic security is wiretap channel. In this section we study the Gaussian MIMOME wiretap channel where a transmitter wishes to communicate a confidential message to a legitimate receiver in the presence of eavesdroppers, while the eavesdroppers should not be able to decode the confidential message. Each node in the network is equipped with arbitrary number of antennas. Furthermore, channels are time varying, and there is no channel state information available at the transmitter (CSIT) with respect to eavesdroppers' channels; and transmitter only has access to delayed CSIT of the channel to the legitimate receiver.

We first describe the system model, and then present the complete characterization of secure degrees of freedom (SDoF) of blind MIMOME wiretap channel with delayed CSIT for all antenna configurations. We then present the achievable scheme, which strictly improves the state-of-the-art achievable scheme for this network.Finally, we prove the converse. To this aim, we present 4 key inequalities that provide lower bounds on the received signal dimensions at various receivers which supply different types of CSIT. We use these inequalities to prove the converse.

3.2.1 System Model and Main Results

Throughout this section we use small letters (e.g. x) for scalars, arrowed letters (e.g. $\vec{x}$) for vectors, capital letters (e.g. X) for matrices, and calligraphic font (e.g. $\mathcal{X}$) for sets. For any scalar variable x we use $[x]^+$ to denote $max(x, 0)$. Moreover, for a random vector $\vec{x}$, we denote its covariance matrix by $K_{\vec{x}}$. Landau notation $x(n) = o(n)$ is used to denote $\lim_{n\to\infty} \frac{x(n)}{n} = 0$. We use $\det[A]$ to denote determinant of matrix A; and A^H denotes Hermitian transpose of matrix A. I_m denotes the identity matrix of size $m\times m$; and $A\otimes B$ denotes the Kronecker product of matrices A, B. $[x_1; x_2; \ldots; x_n]$ denotes a $n \times 1$ vector with the i-th element being x_i.

We consider the Gaussian multiple-input multiple-output multiple-eavesdropper (MIMOME) wiretap channel depicted in Figure 3.1, which consists of a transmitter (Tx) equipped with m antennas and $k+1$ receivers $\mathrm{Rx}_1, \mathrm{Rx}_2, \ldots, \mathrm{Rx}_{k+1}$, where Rx_i ($i = 1, \ldots, k + 1$) is equipped with n_i antennas ($m, n_1, \ldots, n_{k+1} \in \mathbb{N}$). Throughout this section we denote the maximum of

$n_2, \ldots, n_{k+1}$ by n_{max}, and the corresponding receiver by Rx_{max}. In other words, Rx_{max} is the eavesdropper with the most number of antennas (n_{max} antennas).[2] Tx has a secret message for Rx_1 (legitimate receiver), while $\mathrm{Rx}_2, \ldots, \mathrm{Rx}_{k+1}$ are eavesdroppers.

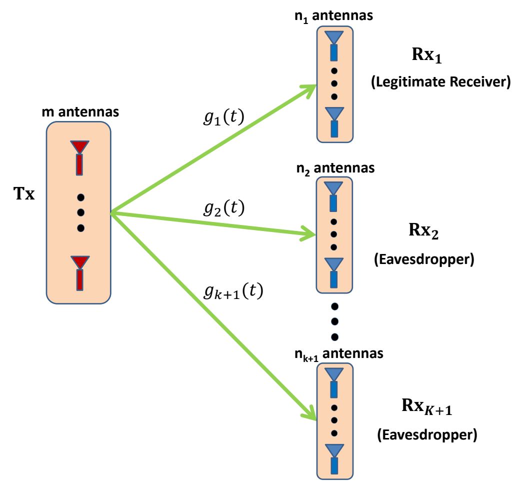


Figure 3.1: Network configuration for the blind MIMOME wiretap channel with k eavesdroppers, where $\mathrm{Rx}_2, \ldots, \mathrm{Rx}_{k+1}$ do not supply any CSIT, and Rx_1 only supplies delayed CSIT.

The received signal at Rx_j ($j \in \{1, \ldots, k+1\}$) at time t is given by

$$\vec{y}_j(t) = g_j(t)\vec{x}(t) + \vec{z}_j(t), \tag{3.1}$$

where $\vec{x}(t) \in \mathbb{C}^m$ is the transmit signal vector of Tx; $g_j(t) \in \mathbb{C}^{n_j \times m}$ indicates the channel matrix from Tx to Rx_j; and $\vec{z}_j(t) \sim \mathcal{CN}(0, I_{n_j})$. The channel coefficients comprising the channel matrix $g_j(t)$ are i.i.d, and also i.i.d. across time, antennas, and receivers, and they are drawn from a continuous distribution, where the absolute value of each element of $g_j(t)$ is bounded by a large number d_{max}. We denote by $\mathcal{G}(t) \triangleq \{g_1(t), \ldots, g_{k+1}(t)\}$ the set of all channel coefficients at time t. In addition, we denote by $\mathcal{G}^n$ the set of all channel coefficients from time 1 to n, i.e.,

$$\mathcal{G}^n \triangleq \{g_j(t) : j \in \{1, \ldots, k+1\}, t \in \{1, \ldots, n\}\}.$$

[2]If there are multiple eavesdroppers with n_{max} antennas, we consider the eavesdropper with smallest index to be Rx_{max}.

Moreover, we use the following notation throughout this section:

$$\vec{x}^t \triangleq \begin{bmatrix} \vec{x}(1) \\ \vdots \\ \vec{x}(t) \end{bmatrix}, \vec{y}_j^t \triangleq \begin{bmatrix} \vec{y}_j(1) \\ \vdots \\ \vec{y}_j(t) \end{bmatrix}, \vec{z}_j^t \triangleq \begin{bmatrix} \vec{z}_j(1) \\ \vdots \\ \vec{z}_j(t) \end{bmatrix}, G_j^t \triangleq diag(G_j(1), \ldots, G_j(t)),$$

where $diag(G_j(1), \ldots, G_j(t))$ is the block diagonal matrix which has $G_j(1), \ldots, G_j(t)$ on its diagonal. Similarly,

$$\vec{x}^{t_0:t_1} \triangleq \begin{bmatrix} \vec{x}(t_0) \\ \vdots \\ \vec{x}(t_1) \end{bmatrix}, \vec{y}_j^{t_0:t_1} \triangleq \begin{bmatrix} \vec{y}_j(t_0) \\ \vdots \\ \vec{y}_j(t_1) \end{bmatrix}, \vec{z}_j^{t_0:t_1} \triangleq \begin{bmatrix} \vec{z}_j(t_0) \\ \vdots \\ \vec{z}_j(t_1) \end{bmatrix}, G_j^{t_0:t_1} \triangleq diag(G_j(t_0), \ldots, G_j(t_1)).$$

The transmitter obeys an average power constraint, $\frac{1}{n}E\{\|\vec{x}^n\|^2\} \leq p$. We assume delayed channel state information at the transmitter (CSIT) with respect to the channel to the legitimate receiver (Rx_1); however, the transmitter does not have knowledge of the channels to the eavesdroppers. In other words, at time t, only G_1^{t-1} is known precisely to the transmitter; and the transmitter only knows a probability distribution for values of the channels to the eavesdroppers, where we denote the maximum value of such distribution by f_{max}, where $f_{max} = o(\log p)$.

Definition 6. *A code for a communication block length of n for the blind MIMOME wiretap channel with delayed CSIT consists of:*

- *A sequence of encoders, $f^{(n)} = (f_1^{(n)}, \ldots, f_n^{(n)})$, where at time t,*

$$\vec{x}(t) = f_t^{(n)}(W, G_1{}^{t-1}),$$

  where message W is uniformly distributed over $\{1, 2, \ldots, |\mathcal{W}(n)|\}$.

- *The corresponding decoder $F^{(n)}$ at Rx_1, where*

$$\hat{W} = F^{(n)}(\vec{y}_1^n, \mathcal{G}^n).$$

- *The error probability of communication is defined to be*

$$p_e^{(n)} = \Pr(W \neq \hat{W}).$$

Based on the above definition, we now define the secure degrees of freedom (SDoF) of the blind MIMOME wiretap channel with delayed CSIT.

Definition 7. *d secure degrees of freedom are achievable if there exists a sequence of encoders and decoders $\{f^{(n)}, F^{(n)}\}_{n=1}^{\infty}$, such that*

$$\limsup_{n\to\infty} p_e^{(n)} = 0, \tag{3.2}$$

and

$$\liminf_{p\to\infty} \liminf_{n\to\infty} \frac{\log |\mathcal{W}(n)|}{n \log p} \geq d, \tag{3.3}$$

and Equivocation condition is satisfied:

$$\limsup_{p\to\infty} \limsup_{n\to\infty} \frac{I(W; \vec{y}_j^n)}{n \log p} = 0, \quad 2 \leq j \leq k+1. \tag{3.4}$$

*We define $\mathcal{D}$ to be the convex closure of the set of all achievable d's. We also define secure degrees of freedom (*SDoF*) to be the supremum of all $d \in \mathcal{D}$.*

Remark 6. *Equivocation condition in (3.4) means that the prelog factor of the equivocation rate to eavesdroppers should vanish as $n \to \infty$. The Equivocation condition (3.4) is weaker than the condition $\limsup_{n\to\infty} \frac{I(W;\vec{y}_j^n)}{n} = 0, \quad 2 \leq j \leq k+1$, considered in some prior works including [68, 19]. However, one can combine our achievable schemes presented in Section 3.2.2 for blind MIMOME wiretap channel with delayed CSIT with random binning [90] to satisfy the latter condition as well.*

For the problem of blind MIMOME wiretap channel with delayed CSIT, Yang et al. [95] have characterized SDoF for the special case where one of the receivers has more antennas than the transmitter. However, SDoF is in general unknown, and [95] only provides lower bounds on SDoF for the general case.

The main result of this section is the following theorem, which completely characterizes SDoF for blind MIMOME wiretap channel with delayed CSIT for all antenna configurations by improving the best known achievable schemes and providing tight upper bounds.

Theorem 4. *For the blind MIMOME wiretap channel with delayed CSIT, let* $\bar{m} = min(m, n_1 + n_{max})$ *and* $\bar{n} = min(n_1, n_{max})$*. Then,* SDoF *is characterized as follows:*

$$\text{SDoF} = \begin{cases} [m - n_{max}]^+ & \textit{if } m \le max(n_1, n_{max}) \\ \frac{n_1(\bar{m} - n_{max})}{\bar{m} - n_{max} + \bar{n}} & \textit{if } m > max(n_1, n_{max}). \end{cases} \tag{3.5}$$

Note that the SDoF of blind MIMOME wiretap channel with delayed CSIT does not decrease with increasing the number of eavesdroppers; rather, SDoF is only a function of n_{max}, the maximum number of antennas on a single eavesdropper. As a special case, Theorem 4 implies that the achievable scheme presented in [94] for blind MISO wiretap channel with delayed CSIT, which achieves $\frac{1}{2}$, is indeed optimal.[3]

In order to better understand how the result compares with prior works from the achievability perspective, Table 3.1 presents two classes of antenna configurations for which our results strictly improve the existing achievable schemes. In particular, the table provides an example for each case, and states the achieved secure degrees of freedom by both [95] and Theorem 4. Thus, Theorem 4 strictly improves the existing achievable schemes in cases such as

[3]This special case has also been independently studied in [71].

Antenna Configuration	Example	Achieved SDoF in [95]	SDoF
$n_1 \leq n_{max} < m$, $m \leq n_1 + n_{max}$	$m = 4, n_1 = 2, n_{max} = 3$, illustrated in Fig. 3.2	$\frac{n_1(m-n_{max})}{m}$	$\frac{n_1(m-n_{max})}{m-n_{max}+n_1}$
$n_1 \leq n_{max}$, $m > n_1 + n_{max}$	$m = 3, n_1 = 1, n_{max} = 2$, illustrated in Fig. 3.3	$\frac{n_1^2}{n_1+n_{max}}$	$\frac{n_1}{2}$

Table 3.1: Comparison of achievability results for 2 different classes of antenna configuration

$n_1 \leq n_{max} < m \leq n_1 + n_{max}$, and $m \geq n_1 + n_{max}$, where $n_{max} > n_1$. Moreover, as we will see in Section 3.2.2, we provide a single unified achievable scheme for all antenna configurations which satisfy $max(n_1, n_{max}) < m$.

In the following sections we provide the proof for Theorem 4, and explain the key ideas behind the proof.

3.2.2 Proof of Achievability

In this section we present the achievable schemes for all antenna configurations. At a high level, our scheme transmits two types of symbols: information symbols, which together constitute the confidential message, and artificial noise symbols. By using an appropriate linear precoder and utilizing delayed CSIT supplied by the legitimate receiver, we align artificial noise symbols into a smaller linear subspace at the legitimate receiver so that some room is left for decoding information symbols, while allowing artificial noise to occupy the whole received signal space at the eavesdroppers. This completely drowns information symbols in artificial noise at the eavesdroppers so that eavesdroppers will not be able to decode them, while it allows the legitimate receiver to successfully decode the information symbols.

Throughout the presentation of the achievable schemes, for simplicity and without loss of generality we assume that $n = n'b$, where b is the block length of communication for our scheme, and n' is the number of blocks. In fact, we repeat the same transmission scheme for all blocks. We first present the achievable scheme for the case of $m \leq max(n_1, n_{max})$.[4]

Case of $m \leq max(n_1, n_{max})$

Note that for the case where $m \leq n_{max}$, Theorem 4 suggests that SDoF = 0; so there is nothing to prove on the achievability side. Thus, let us consider the case where $n_{max} < m \leq n_1$. In this case we will show that $d = m - n_{max}$ secure degrees of freedom is achievable. In other words, we show how to securely deliver $m - n_{max}$ information symbols to Rx_1 in each time slot ($b = 1$). In particular, in every time slot, each of the first n_{max} transmit antennas sends a distinct artificial noise symbol, while each of the antennas with index $n_{max} + 1, \ldots, m - 1, m$ sends a distinct new information symbol. Consequently, Rx_1 recovers all symbols almost surely, including the $m - n_{max}$ information symbols, since it receives n_1 equations in m unknowns where $m \leq n_1$, hence satisfying (3.3). By using an appropriate code for the Gaussian MIMO channel between Tx and Rx_1, when the block length of communication goes to infinity (i.e. $n \to \infty$), the error of communication goes to

[4]The achievable scheme for the case $m \leq max(n_1, n_{max})$ is presented in [95]; however, we state it here for completion and because its analysis serves as an introduction to analysis of the case $m > max(n_1, n_{max})$.

zero, satisfying (3.2).

Moreover, the eavesdroppers cannot decode any of the information symbols because Rx_{max} essentially receives n_{max} equations regarding n_{max} artificial noise symbols (undesired symbols) and $(m - n_{max})$ information symbols. Hence, the information symbols are completely drowned in artificial noise and Equivocation condition (3.4) is satisfied.

After providing the intuitive reason why the achievable scheme satisfies conditions (3.3)-(3.4), we now rigorously prove that under the proposed scheme, conditions (3.3)-(3.4) are satisfied. To this aim, we state a lemma that will be useful in the proof of achievability.

Lemma 7. *For a fixed matrix A,*

$$\lim_{p \to \infty} \frac{\log \det[I + pAA^H]}{\log p} = rank[A]. \tag{3.6}$$

Proof of Lemma 7 follows from straightforward linear algebra and can be found in [95].

We first specify the transmit signals, and then use them to show that conditions (3.3)-(3.4) are satisfied. At time slot t, $\vec{u}_{n_{max} \times 1} \in \mathcal{CN}(0, \frac{p}{m} I_{n_{max}})$ denotes the vector of artificial noise symbols, which are transmitted by antennas $1, \ldots, n_{max}$, and the vector of information symbols by $\vec{v}_{(m-n_{max}) \times 1} \in \mathcal{CN}(0, \frac{p}{m} I_{(m-n_{max})})$, which are transmitted by antennas $n_{max} + 1, \ldots, m$. As a result,

$$\vec{x} = \begin{bmatrix} \vec{u}_{n_{max} \times 1} \\ \\ \vec{v}_{(m-n_{max}) \times 1} \end{bmatrix}, \qquad K_{\vec{x}} = \frac{p}{m} I_m, \qquad K_{\vec{x}|\vec{v}} = \frac{p}{m} \begin{bmatrix} I_{n_{max}} & 0 \\ \\ 0 & 0 \end{bmatrix}. \tag{3.7}$$

Therefore, for any $j \in \{1, 2, \ldots, k+1\}$,

$$
\begin{aligned}
&\lim_{p\to\infty} \frac{I(\vec{v};\vec{y}_j|\mathcal{G})}{\log p} = \lim_{p\to\infty} E_{\mathcal{G}}\{\frac{h(\vec{y}_j|\mathcal{G}) - h(\vec{y}_j|\vec{v},\mathcal{G})}{\log p}\} \\
&= \lim_{p\to\infty} E_{\mathcal{G}}\{\frac{\log\det[I + G_j K_{\vec{x}} G_j^H] - \log\det[I + G_j K_{\vec{x}|\vec{v}} G_j^H]}{\log p}\} \\
&\overset{(a)}{=} E_{\mathcal{G}}\{\lim_{p\to\infty} \frac{\log\det[I + G_j K_{\vec{x}} G_j^H] - \log\det[I + G_j K_{\vec{x}|\vec{v}} G_j^H]}{\log p}\} \\
&= E_{\mathcal{G}}\{\lim_{p\to\infty} \frac{\log\det[I + \frac{p}{m} G_j G_j^H]}{\log p} - \lim_{p\to\infty} \frac{\log\det[I + \frac{p}{m} G_j \begin{bmatrix} I_{n_{max}} & 0 \\ 0 & 0_{(m-n_{max})} \end{bmatrix} G_j^H]}{\log p}\} \\
&\overset{\text{(Lemma 7)}}{=} E_{\mathcal{G}}\{\mathrm{rank}[G_j] - \mathrm{rank}[G_j \begin{bmatrix} I_{n_{max}} & 0 \\ 0 & 0_{(m-n_{max})} \end{bmatrix}]\},
\end{aligned} \tag{3.8}
$$

where (a) is due to Dominated Convergence Theorem. Furthermore, note that for a random channel realization $\mathcal{G}$,

$$
\mathrm{rank}[G_j] \overset{a.s.}{=} min(n_j, m), \qquad \mathrm{rank}[G_j \begin{bmatrix} I_{n_{max}} & 0 \\ 0 & 0_{(m-n_{max})} \end{bmatrix}] \overset{a.s.}{=} min(n_j, n_{max}). \tag{3.9}
$$

Hence, since we are considering the case where $n_{max} < m \le n_1$, by (3.8), (3.9), we obtain

$$
\lim_{p\to\infty} \frac{I(\vec{v};\vec{y}_1|\mathcal{G})}{\log p} = m - n_{max}, \tag{3.10}
$$

$$
\lim_{p\to\infty} \frac{I(\vec{v};\vec{y}_j|\mathcal{G})}{\log p} = 0, \qquad j = 2, \ldots, k+1, \tag{3.11}
$$

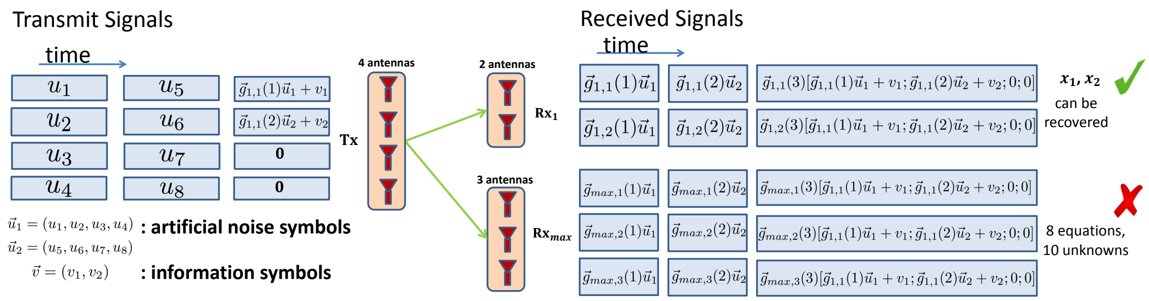


Figure 3.2: The achievable scheme for a simple network configuration that belongs to case where $n_1 \leq n_{max}, m \geq n_1 + n_{max}$. The scheme delivers 2 symbol securely over 3 time slots, achieving SDoF of $\frac{2}{3}$.

which prove that conditions (3.3),(3.4) are satisfied. Hence, the proposed achievable scheme achieves $[m - n_{max}]^+$ secure degrees of freedom when $m \leq max(n_1, n_{max})$.

Case of $m > max(n_1, n_{max})$

In this case the scheme will securely deliver $n_1(\bar{m} - n_{max})$ information symbols over $\bar{m} - n_{max} + \bar{n}$ time slots ($b = \bar{m} - n_{max} + \bar{n}$). The scheme is presented in two phases. The first phase takes $\bar{n}$ time slots, during which only artificial noise symbols are transmitted. Then, during the second phase, which takes $\bar{m} - n_{max}$ time slots, some of artificial noise equations are retransmitted together with information symbols in such a way that they completely mask the information symbols at the eavesdroppers, while the information symbols can be recovered at the legitimate receiver. Since throughout the description of our achievable scheme and its analysis only $\bar{m}$ transmit antennas are used at any point in time, we only focus on the first $\bar{m}$ transmit antennas for the sake of simplicity and ignore the rest. Further, we implicitly consider the proper scaling which is needed to satisfy the power constraint. We now present the details of our scheme.

Phase 1: This phase takes $\bar{n}$ time slots. At $t = 1, 2, \ldots, \bar{n}$, each of the $\bar{m}$ transmit antennas sends a new artificial noise symbol. Thus, since channel coefficients are i.i.d. and drawn from a continuous distribution, what Rx_1 receives on its first $(\bar{m} - n_{max})$ antennas are almost surely linearly independent of equations received by antennas of Rx_{max}, hence not recoverable by Rx_{max}. Similar result holds for other eavesdroppers. Hence, by the end of phase 1, Rx_1 obtains $\bar{n}(\bar{m} - n_{max})$ linearly independent noise equations that are not recoverable by eavesdroppers.

Phase 2: This phase takes $\bar{m} - n_{max}$ time slots. In each time slot $t \in \{\bar{n}+1, \ldots, \bar{m} - n_{max} + \bar{n}\}$, transmit signals by the $\bar{m}$ transmit antennas are as follows:

$$\vec{x}(t) = \begin{bmatrix} \vec{u}'_{\bar{n}}(t) \\ \\ \vec{0}_{\bar{m}-\bar{n}} \end{bmatrix} + \begin{bmatrix} \vec{v}_{n_1}(t) \\ \\ \vec{0}_{\bar{m}-n_1} \end{bmatrix}, \tag{3.12}$$

where $\vec{u}'_{\bar{n}}(t)$ is a vector comprised of $\bar{n}$ linearly independent artificial noise equations known by Rx_1 (ignoring AWGN), which are not recoverable by eavesdroppers, and are produced as the result of Phase 1 of the scheme. Note that these artificial noise equations are known to the transmitter by the end of Phase 1 because it has access to delayed CSIT of the legitimate receiver as well as all artificial noise symbols. We will refer to these artificial noise equations as artificial noise symbols for simplicity. Moreover, $\vec{v}_{n_1}(t)$is the vector of information symbols.

As a result of this transmission scheme, in each time slot of Phase 2, Rx_1 cancels the transmitted artificial noise symbols from its received signal, and recovers the n_1 desirable information symbols. Therefore, Rx_1 recovers $n_1(\bar{m} - n_{max})$ information symbols in total, satisfying (3.3).

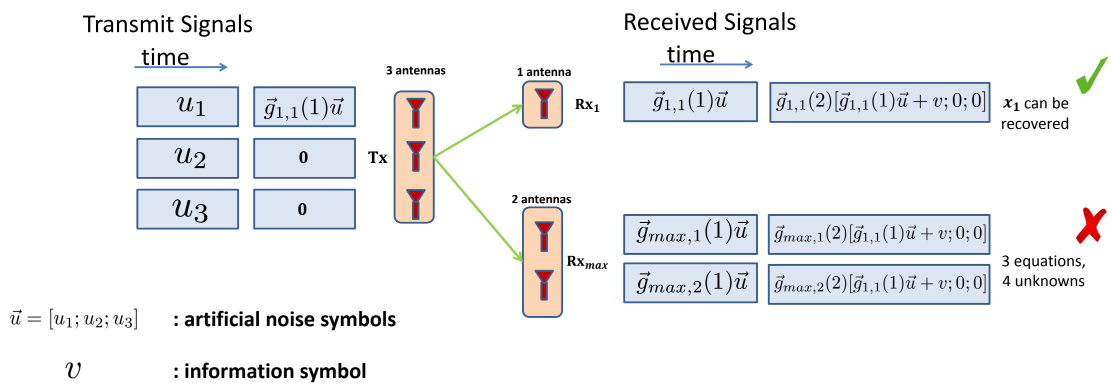


Figure 3.3: The achievable scheme for a simple network configuration that belongs to case where $n_1 \leq n_{max}, m \geq n_1 + n_{max}$. The scheme delivers 1 symbol securely over 2 timeslots, achieving SDoF of $\frac{1}{2}$.

In order to explain why Equivocation condition (3.4) is satisfied , we focus on Rx_{max} as the argument is similar for other eavesdroppers. We consider the two antenna configurations $\bar{n} = n_1 \leq n_{max}$, and $\bar{n} = n_{max} \leq n_1$. In the case of $\bar{n} = n_1 \leq n_{max}$, each active antenna is sending one information symbol plus an artificial noise symbol which is not known to Rx_{max}. Therefore, Rx_{max} cannot recover any of the information symbols. Moreover, in the case of $\bar{n} = n_{max} \leq n_1$, the transmitter is sending n_{max} artificial noise symbols in each time slot which are not known to Rx_{max}. Thus, since Rx_{max} has n_{max} antennas, the transmitted artificial noise symbols span the entire space of received signals at Rx_{max} and do not allow for any information symbol to be recovered, satisfying (3.4).

Hence, overall $n_1(\bar{m} - n_{max})$ information symbols are delivered securely to Rx_1 over $\bar{m} - n_{max} + \bar{n}$ time slots, and the scheme achieves $\frac{n_1(\bar{m}-n_{max})}{\bar{m}-n_{max}+\bar{n}}$ secure degrees of freedom. Thus, the proposed achievable scheme strictly improves the existing achievable schemes [95] for the case $n_1 < n_{max} < m \leq n_1 + n_{max}$, illustrated in Figure 3.2, and $m \geq n_1 + n_{max}$, where $n_{max} > n_1$, as illustrated in Figure 3.3.

After explaining why the proposed achievable scheme achieves $\frac{n_1(\bar{m}-n_{max})}{\bar{m}-n_{max}+\bar{n}}$ secure degrees of freedom, we now prove that the conditions (3.3),(3.4) are indeed satisfied. Let $\vec{u} \sim \mathcal{CN}(0, \frac{p}{\bar{m}} I_{\bar{n}\bar{m}}), \vec{v} \sim \mathcal{CN}(0, \frac{p}{\bar{m}} I_{n_1(\bar{m}-n_{max})})$. The transmit signal is as follows:

$$\vec{x}^{\bar{m}-n_{max}+\bar{n}} = \begin{bmatrix} I_{\bar{n}\bar{m}} & 0_{\bar{n}\bar{m}\times n_1(\bar{m}-n_{max})} \\ A & I_{\bar{m}-n_{max}} \otimes \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0_{\bar{m}-n_1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \vec{u} \\ \vec{v} \end{bmatrix}, \tag{3.13}$$

where

$$A \triangleq \begin{bmatrix} A(\bar{n}+1) \\ \vdots \\ A(\bar{m}-n_{max}+\bar{n}) \end{bmatrix}, \quad \text{and} \quad A(t) \triangleq \begin{bmatrix} \vec{g}_{1,t-\bar{n}}(1) & 0 & \dots & 0 \\ 0 & \vec{g}_{1,t-\bar{n}}(2) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \vec{g}_{1,t-\bar{n}}(\bar{n}) \\ & 0_{(\bar{m}-\bar{n})\times\bar{n}\bar{m}} & & \end{bmatrix}, \tag{3.14}$$

where $\vec{g}_{1,i}(t)$ is the channel vector comprised of coefficients of channels between transmitter and the i-th receive antenna of Rx_1 at time t . Hence, the received signal at Rx_j ($j = 1, \ldots, k+1$) is

$$\vec{y}_j^{\bar{m}-n_{max}+\bar{n}} = \begin{bmatrix} G_j^{\bar{n}} & 0_{\bar{n}n_j \times n_1(\bar{m}-n_{max})} \\ G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} A & G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} I_{\bar{m}-n_{max}} \otimes \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0_{\bar{m}-n_1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \vec{u} \\ \vec{v} \end{bmatrix} + \vec{z}_j^{\bar{m}-n_{max}+\bar{n}}. \tag{3.15}$$

We now show that (3.3)-(3.4) are satisfied. Using similar steps as in (3.8), we obtain

$$\begin{aligned} &\lim_{p\to\infty} \frac{I(\vec{v}; \vec{y}_j^{\bar{m}-n_{max}+\bar{n}})}{\log p} \\ &= E_{\mathcal{G}^{\bar{m}-n_{max}+\bar{n}}}\{\text{rank} \begin{bmatrix} G_j^{\bar{n}} & 0_{\bar{n}n_j \times n_1(\bar{m}-n_{max})} \\ G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} A & G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} I_{\bar{m}-n_{max}} \otimes \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0_{\bar{m}-n_1} \end{bmatrix} \end{bmatrix}\} \\ &\quad - E_{\mathcal{G}^{\bar{m}-n_{max}+\bar{n}}}\{\text{rank} \begin{bmatrix} G_j^{\bar{n}} \\ G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} A \end{bmatrix}\}. \end{aligned} \tag{3.16}$$

Due to the structure of A,

$$\text{rank}\begin{bmatrix} G_j^{\bar{n}} \\ \\ G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}}A \end{bmatrix} \overset{a.s.}{=} \begin{cases} \text{rank}[G_j^{\bar{n}}], & \text{for } j=1 \\ \\ \text{rank}[G_j^{\bar{n}}] + \text{rank}[G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}}A], & \text{for } j>1 \end{cases} \tag{3.17}$$

and

$$\text{rank}\begin{bmatrix} G_j^{\bar{n}} & 0_{\bar{n}n_j \times n_1(\bar{m}-n_{max})} \\ \\ G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}}A & G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} I_{\bar{m}-n_{max}} \otimes \begin{bmatrix} I_{n_1} & 0 \\ \\ 0 & 0_{\bar{m}-n_1} \end{bmatrix} \end{bmatrix}$$

$$\overset{a.s.}{=} \begin{cases} \text{rank}[G_j^{\bar{n}}] + \text{rank}\begin{bmatrix} G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}} I_{\bar{m}-n_{max}} \otimes \begin{bmatrix} I_{n_1} & 0 \\ \\ 0 & 0_{\bar{m}-n_1} \end{bmatrix} \end{bmatrix}, & \text{for } j=1 \\ \\ \text{rank}[G_j^{\bar{n}}] + \text{rank}[G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}}A], & \text{for } j>1 \end{cases}$$

$$\overset{a.s.}{=} \begin{cases} \text{rank}[G_j^{\bar{n}}] + n_1(\bar{m}-n_{max}), & \text{for } j=1 \\ \\ \text{rank}[G_j^{\bar{n}}] + \text{rank}[G_j^{\bar{n}+1:\bar{m}-n_{max}+\bar{n}}A], & \text{for } j>1. \end{cases} \tag{3.18}$$

Using (3.16)-(3.18), one can readily see that conditions (3.3)-(3.4) are satisfied. Hence, the proposed achievable scheme achieves $\frac{n_1(\bar{m}-n_{max})}{\bar{m}-n_{max}+\bar{n}}$ secure degrees of freedom when $m > max(n_1, n_{max})$; and thus, achievability proof is complete.

3.2.3 Proof of Converse

In this section we present the converse proof for Theorem 4. Note that for any antenna configuration $(m, n_1, n_2, \ldots, n_{k+1})$, if some of the eavesdroppers are removed from the network, SDoF will not decrease; and this is due to removing some of the Equivocation constraints on maximizing the secure rate. Hence, to prove the converse we first remove all the eavesdroppers except Rx_{max} from the network. We start by proving the converse for the case where $m \leq max(n_1, n_{max})$.[5]

Proof of Converse for $m \leq max(n_1, n_{max})$

We first state a lemma that will be used throughout this section.

Lemma 8. *Consider two receivers Rx_1, Rx_2 with n_1, n_2 antennas, where Rx_2 supplies delayed CSIT or no CSIT. For any fixed n, and any encoding strategy $f^{(n)}$,*

$$\frac{h(\vec{y}_2^n|\mathcal{G}^n)}{min(m, n_2)} + n \times o(\log p) \geq \frac{h(\vec{y}_1^n, \vec{y}_2^n|\mathcal{G}^n)}{min(m, n_1 + n_2)}.$$

Lemma 8 provides a lower bound on the received signal dimension at a receiver which does not supply perfect CSIT. In other words, it relates the received signal dimension of a receiver which supplies delayed or no CSIT to that of another receiver. Proof of Lemma 8 follows from channel symmetry; and it can be found in [86, 95].

We now prove the converse for $m \leq max(n_1, n_{max})$. Suppose d secure DoF is achievable. Therefore, by Fano's inequality,

$$n(R - \epsilon_n) \leq I(W; \vec{y}_1^n|\mathcal{G}^n) \leq I(W; \vec{y}_1^n, \vec{y}_{max}^n|\mathcal{G}^n)$$

[5]Proof of converse for the case when $m \leq max(n_1, n_{max})$ has been presented in [95]; nevertheless, we provide the proof here for completion.

$$
\begin{aligned}
&\overset{(a)}{\leq} h(\vec{y}_1^n, \vec{y}_{max}^n | \mathcal{G}^n) - h(\vec{y}_{max}^n | W, \mathcal{G}^n) \\
&= h(\vec{y}_1^n, \vec{y}_{max}^n | \mathcal{G}^n) - h(\vec{y}_{max}^n | \mathcal{G}^n) + I(W; \vec{y}_{max}^n | \mathcal{G}^n) \\
&\overset{(3.4)}{\leq} h(\vec{y}_1^n, \vec{y}_{max}^n | \mathcal{G}^n) - h(\vec{y}_{max}^n | \mathcal{G}^n) + n \times o(\log p) \\
&\overset{(\text{Lemma 8})}{\leq} \frac{m - min(m, n_{max})}{min(m, n_{max})} \times h(\vec{y}_{max}^n | \mathcal{G}^n) + n \times o(\log p) \\
&= \frac{[m - n_{max}]^+}{n_{max}} \times h(\vec{y}_{max}^n | \mathcal{G}^n) + n \times o(\log p) \\
&\leq \frac{[m - n_{max}]^+}{n_{max}} \times n n_{max} \log p + n \times o(\log p) \\
&= [m - n_{max}]^+ \times n \log p + n \times o(\log p),
\end{aligned}
$$

where (a) holds since $h(\vec{y}_1^n | \vec{y}_{max}^n, W, \mathcal{G}^n) \geq h(\vec{y}_1^n | \vec{x}^n, \vec{y}_{max}^n, W, \mathcal{G}^n) = h(\vec{z}_1^n) > 0$. Hence, by dividing both sides of the above inequality by $n \log p$ and taking the limit $n \to \infty$ and $p \to \infty$, the result follows.

Proof of Converse for $m > max(n_1, n_{max})$

Before presenting the converse proof for $m > max(n_1, n_{max})$, we consider a notation that will be used throughout the proof. Let $\text{Rx}_{max,1}$ denote the first $\bar{n}$ antennas on Rx_{max} with the corresponding received signal of $\vec{y}_{max,1}^n$ over n time slots. Further, let $\text{Rx}_{max,2}$ denote the remaining $n_{max} - \bar{n}$ antennas on Rx_{max} with the corresponding received signal of $\vec{y}_{max,2}^n$. Finally, we denote by $\text{Rx}_{1,1}$ the first $\bar{m} - n_{max}$ antennas on Rx_1 with the corresponding received signal of $\vec{y}_{1,1}^n$.

We first present the tools that are used to prove the converse for this case. The first lemma, Least Alignment Lemma, implies that once the transmitter(s) in a network has no CSIT with respect to a certain receiver, the least amount of alignment will occur at that receiver, meaning that transmit signals will occupy the maximal signal dimensions at that receiver. As a result, for the specific case of having two n_0-antenna receivers $\mathrm{Rx}_1, \mathrm{Rx}_2$, if the transmitter does not have access to any CSIT with respect to Rx_2, then the received signal dimension at Rx_2 will be greater than Rx_1.

Lemma 9. (Least Alignment Lemma) *Consider two receivers* $\mathrm{Rx}_1, \mathrm{Rx}_2$ *with* n_0 *antennas, where* Rx_2 *supplies no CSIT. Then, for a given* $n \in \mathbb{N}$ *and any encoding strategy* $f^{(n)}$ *as defined in Definition 6,*

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq h(\vec{y}_2^n|\mathcal{G}^n) + n \times o(\log p).$$

Remark 7. *Lemma 9 holds irrespective of the type of CSIT supplied by* Rx_1. *Furthermore, Lemma 9 holds for arbitrary number of transmitters with arbitrary number of transmit antennas; this can be shown via following the same steps as in the proof presented in Appendix B.1. Therefore, Least Alignment Lemma is a general inequality that can be applied to lower bound the received signal dimension at any receiver which supplies no CSIT. In fact, Least Alignment Lemma relates the received signal dimension at a receiver which supplies no CSIT to the dimension at other receivers with the same number of antennas, and as a result, proves to be an important tool in analyzing networks with heterogeneous CSIT, where some receivers supply no CSIT, while others supply some form of CSIT to the transmitter(s).*

[47] provided a proof of the lemma for single-antenna receivers, which was limited to networks whose transmitter(s) were only able to employ linear encoding schemes.[6] However, Davoodi and Jafar [21] provided the first proof of the inequality in Lemma 9 (for single-antenna receivers) under general encoding schemes. Their proof was based on a novel analysis of the Aligned Image Sets at receivers which supply imperfect CSIT. The proof was used to settle important conjectures regarding networks with imperfect CSIT [21]. Proof of Lemma

[6]Extension of the proof to networks with multiple antenna receivers appeared in [45].

9, presented in Appendix B.1, extends the result to the MIMO setting.

The following lemma relates the received signal dimensions at 2 receivers supplying no CSIT.

Lemma 10. *Consider receivers* $\mathrm{Rx}_1, \mathrm{Rx}_2$ *which supply no CSIT, with* n_1, n_2 *antennas, where* $n_1 \geq n_2$*. Then, for a given* $n \in \mathbb{N}$ *and any encoding strategy* $f^{(n)}$ *as defined in Definition 6,*

$$\frac{h(\vec{y}_1^n|\mathcal{G}^n)}{min(m, n_1)} \leq \frac{h(\vec{y}_2^n|\mathcal{G}^n)}{min(m, n_2)} + n \times o(\log p).$$

Proof of Lemma 10 follows from channel symmetry, and can be found in prior works, including [85].

Proposition 4. *If* $m > max(n_1, n_{max})$*, then,*

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq \frac{n_1}{\bar{n}} h(\vec{y}_{max,1}^n|\mathcal{G}^n) + n \times o(\log p).$$

Proof of Proposition 4 follows from Lemma 9 and Lemma 10, and is postponed to Appendix B.2.

Finally, the following lemma, provides a lower bound on the dimension of joint received signals at a collection of receivers, where some receivers supply no CSIT.

Lemma 11. *Consider receivers* $\mathrm{Rx}_1, \mathrm{Rx}_2, \mathrm{Rx}_3$ *with* n_1, n_2, n_3 *antennas, where* $n_1, n_2, n_3 > 0$*, and* $m \geq n_1 + n_2 + n_3$*. Further, suppose that* $\mathrm{Rx}_2, \mathrm{Rx}_3$ *supply no CSIT. Then, for a given* $n \in \mathbb{N}$ *and any encoding strategy* $f^{(n)}$ *as defined in Definition 6,*

$$\frac{h(\vec{y}_1^n|\vec{y}_2^n, \vec{y}_3^n, \mathcal{G}^n)}{n_1} \leq \frac{h(\vec{y}_2^n|\vec{y}_3^n, \mathcal{G}^n)}{n_2} + n \times o(\log p).$$

Lemma 11 is proved in Appendix B.3.

Proposition 5. *If $n_1 < n_{max} < m$, for a given $n \in \mathbb{N}$ and any encoding strategy $f^{(n)}$ as defined in Definition 6,*

$$\frac{h(\vec{y}^n_{1,1}|\vec{y}^n_{max}, \mathcal{G}^n)}{\bar{m} - n_{max}} \leq \frac{h(\vec{y}^n_{max,2}|\vec{y}^n_{max,1}, \mathcal{G}^n)}{n_{max} - \bar{n}} + n \times o(\log p).$$

Proof of Proposition 5 follows immediately from substituting $(\vec{y}^n_1, \vec{y}^n_2, \vec{y}^n_3)$ by $(\vec{y}^n_{1,1}, \vec{y}^n_{max,2}, \vec{y}^n_{max,1})$ in the statement of Lemma 11.

We now prove the converse for the case where $m > max(n_1, n_{max})$. Throughout the proof, we use the notation $h(\emptyset|\mathcal{G}^n) = 0$ for convenience. Suppose d secure degrees of freedom is achievable. By Fano's inequality,

$$\begin{aligned}
n(R - \epsilon_n) &\leq I(W; \vec{y}^n_1|\mathcal{G}^n) \leq h(\vec{y}^n_1|\mathcal{G}^n) - h(\vec{y}^n_1|W, \mathcal{G}^n) \\
&\overset{(\text{Lemma } 8)}{\leq} h(\vec{y}^n_1|\mathcal{G}^n) - \frac{n_1}{\bar{m}} h(\vec{y}^n_{max}|W, \mathcal{G}^n) \\
&= h(\vec{y}^n_1|\mathcal{G}^n) + \frac{n_1}{\bar{m}} I(W; \vec{y}^n_{max}|\mathcal{G}^n) - \frac{n_1}{\bar{m}} h(\vec{y}^n_{max}|\mathcal{G}^n) \\
&\overset{(3.4)}{\leq} h(\vec{y}^n_1|\mathcal{G}^n) - \frac{n_1}{\bar{m}} h(\vec{y}^n_{max}|\mathcal{G}^n) + n.o(\log p) \\
&\overset{(\text{Proposition } 4)}{\leq} \frac{n_1}{\bar{n}} h(\vec{y}^n_{max,1}|\mathcal{G}^n) - \frac{n_1}{\bar{m}} h(\vec{y}^n_{max}|\mathcal{G}^n) + n.o(\log p) \\
&= n_1 (\frac{\bar{m} - \bar{n}}{\bar{m}\bar{n}}) h(\vec{y}^n_{max,1}|\mathcal{G}^n) - \frac{n_1}{\bar{m}} h(\vec{y}^n_{max,2}|\vec{y}^n_{max,1}, \mathcal{G}^n) + n.o(\log p) \\
&\leq n_1 (\frac{\bar{m} - \bar{n}}{\bar{m}}) n \log p - \frac{n_1}{\bar{m}} h(\vec{y}^n_{max,2}|\vec{y}^n_{max,1}, \mathcal{G}^n) + n.o(\log p). \qquad (3.19)
\end{aligned}$$

We now study the two cases $n_{max} \leq n_1$, and $n_{max} > n_1$. If $n_{max} \leq n_1$, then $\frac{n_1}{\bar{m}} h(\vec{y}^n_{max,2}|\vec{y}^n_{max,1}, \mathcal{G}^n)$, which is the second term on the RHS of (3.19) will equal zero since $\text{Rx}_{max,2}$ is an empty set of antennas when $n_{max} \leq n_1$. As a result, since $\bar{n} = n_{max}$,

$$n(R - \epsilon_n) \leq n_1 (\frac{\bar{m} - n_{max}}{\bar{m} - n_{max} + \bar{n}}) n \log p + n.o(\log p),$$

where by dividing both sides of the inequality by $n \log p$ and taking the limit $n \to \infty$, $p \to \infty$, the converse proof is obtained for the case where $n_{max} \leq n_1$.

We now consider the case where $n_{max} > n_1$. For this case we derive a second bound on the secure rate, and then merge it with (3.19) to obtain the converse proof. Again, by Fano's inequality we obtain

$$\begin{aligned} n(R-\epsilon_n) &\leq I(W;\vec{y}_1^n|\mathcal{G}^n) \leq I(W;\vec{y}_1^n,\vec{y}_{max}^n|\mathcal{G}^n) \leq h(\vec{y}_1^n,\vec{y}_{max}^n|\mathcal{G}^n) - h(\vec{y}_{max}^n|W,\mathcal{G}^n) \\ &\overset{(3.4)}{=} h(\vec{y}_1^n,\vec{y}_{max}^n|\mathcal{G}^n) - h(\vec{y}_{max}^n|\mathcal{G}^n) + n.o(\log p) \\ &\overset{(a)}{=} h(\vec{y}_{1,1}^n,\vec{y}_{max}^n|\mathcal{G}^n) - h(\vec{y}_{max}^n|\mathcal{G}^n) + n.o(\log p) = h(\vec{y}_{1,1}^n|\vec{y}_{max}^n,\mathcal{G}^n) + n.o(\log p) \\ &\overset{\text{(Proposition 5)}}{\leq} (\frac{\bar{m}-n_{max}}{n_{max}-\bar{n}})h(\vec{y}_{max,2}^n|\vec{y}_{max,1}^n,\mathcal{G}^n) + n.o(\log p), \end{aligned} \quad (3.20)$$

where (a) holds since either $m > n_1 + n_{max}$, in which case the equality is obvious as $\vec{y}_{1,1}^n = \vec{y}_1^n$, or $m \leq n_1 + n_{max}$, in which case given $(\vec{y}_{1,1}^n, \vec{y}_{max}^n, \mathcal{G}^n)$, one can reconstruct the transmit signals within noise level; and as a result, $h(\vec{y}_1^n|\vec{y}_{1,1}^n,\vec{y}_{max}^n,\mathcal{G}^n) = n.o(\log p)$.

We now linearly combine the two inequalities (3.19), (3.20) and use the fact that $\bar{n} = n_1$ in this case. By multiplying both sides of (3.19) by $\frac{\bar{m}(\bar{m}-n_{max})}{(\bar{m}-\bar{n})(\bar{m}-n_{max}+\bar{n})}$, and multiplying both sides of (3.20) by $\frac{n_1(n_{max}-\bar{n})}{(\bar{m}-\bar{n})(\bar{m}-n_{max}+\bar{n})}$, and then adding the two inequalities together, we obtain the following inequality by considering the assumption that $\bar{n} = n_1$:

$$n(R-\epsilon_n) \leq \frac{n_1(\bar{m}-n_{max})}{\bar{m}-n_{max}+\bar{n}} n \log p + n.o(\log p),$$

where by dividing both sides of the inequality by $n \log p$ and taking the limit $n \to \infty, p \to \infty$, the converse proof is obtained for the case where $n_{max} > n_1$. Hence, the proof of converse is complete, which completes the proof of Theorem 4.

In the next section we study the blind cooperative SISO wiretap channel with delayed CSIT, and characterize its SDoF under linear encoding schemes.

3.3 Blind Cooperative SISO Wiretap Channel with Delayed CSIT

In this section we continue our study of impact of CSIT on secure communication over wireless networks. In particular, we consider the Gaussian single-input single-output (SISO) wiretap channel, where the single-antenna transmitter is *blind* with respect to the state of channels to eavesdroppers, and only has access to *delayed* channel state information (CSIT) of the legitimate receiver, as considered in the previous section. However, in this problem the secure communication is aided via a distributed single-antenna jammer, which does not necessarily have access to the confidential message, but can help the transmitter by jamming the confidential message at the eavesdroppers so that the eavesdroppers would not be able to decode it.

We refer to this problem as "blind cooperative SISO wiretap channel with delayed CSIT". We first describe the linear model setup, in which transmitters are only allowed to transmit linear combinations of symbols that are available to them. We then present the main result, which is complete characterization of the secure Degrees of Freedom (SDoF) when linear coding strategies are employed at the transmitters. Finally, we provide the achievable scheme and converse proof.

3.3.1 System Model and Main Results

We consider the Gaussian wiretap channel depicted in Fig. 3.4, which consists of a transmitter (Tx_1), a jammer (Tx_2), and $k+1$ receivers, where Tx_1 has a secret message for Rx_1 (legitimate receiver), and $Rx_2, \ldots, Rx_{k+1}$ are eavesdroppers. The role of Tx_2, although it does not have access to the secret message[7], is to help Tx_1 communicate its message securely to Rx_1, while $Rx_2, \ldots, Rx_{k+1}$ cannot decode any part of that message. Each node in the network is equipped with a single antenna.

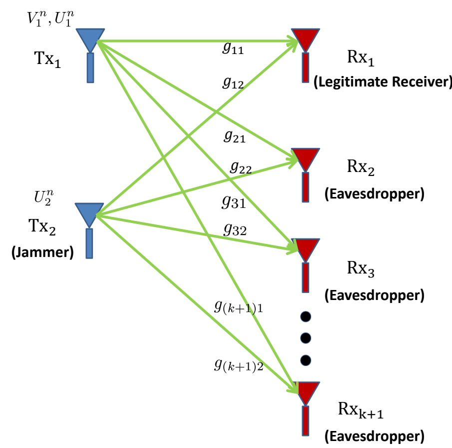


Figure 3.4: Network configuration for the blind cooperative wiretap channel with one jammer and k eavesdroppers.

The received signal at Rx_j ($j \in \{1, \ldots, k+1\}$) at time t is given by

$$\mathbf{y}_j(t) = \mathbf{g}_{j1}(t)\mathbf{x}_1(t) + \mathbf{g}_{j2}(t)\mathbf{x}_2(t) + \mathbf{z}_j(t), \tag{3.21}$$

where $\mathbf{x}_i(t)$ is the transmit signal of Tx_i; $\mathbf{g}_{ji}(t) \in \mathbb{C}$ indicates the channel from Tx_i to Rx_j; and $\mathbf{z}_j(t) \sim \mathcal{CN}(0, 1)$. The channel coefficients $\mathbf{g}_{ji}(t)$ are i.i.d across time and users, and they are drawn from a continuous distribution.[8] We denote by $\mathcal{G}(t)$ the set of all channel coefficients at time t. In addition, we denote by $\mathcal{G}^n$ the set of all channel coefficients from time 1 to n, i.e.,

$$\mathcal{G}^n \triangleq \{\mathbf{g}_{ji}(t) : j \in \{1, \ldots, k+1\}, i \in \{1, 2\}, t \in \{1, \ldots, n\}\}.$$

[7]This assumption is not necessary; and even if the jammer had access to the secret message, the analysis would remain the same.

[8]One can show that the i.i.d. assumption on the channel coefficients can be relaxed, and it is sufficient to assume that the distribution of any channel, conditioned on other channel coefficients, is continuous.

Denoting the vector of transmit signals for Tx_i in a block of length n by $\vec{\mathbf{x}}_i^n$, each transmitter Tx_i obeys an average power constraint, $\frac{1}{n}E\{\|\vec{\mathbf{x}}_i^n\|^2\} \leq P$. We assume delayed channel state information at the transmitters (CSIT) with respect to channels to the legitimate receiver; however, transmitters have no knowledge of eavesdroppers. In other words, at time t, only the states of the past $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{t-1} \triangleq \{\mathbf{g}_{1i}(h) : i = 1, 2, h = 1, \ldots, t-1\}$ are known to the transmitters.

We restrict ourselves to linear coding strategies as defined in [13, 52, 51]. In particular, consider a communication scheme with block length n, in which Tx_1 wishes to communicate a vector $\vec{\mathbf{x}} \in \mathbb{C}^{m_1(n)}$ of $m_1(n) \in \mathbb{N}$ information symbols to Rx_1. Each of the information symbols is a Gaussian random variable with variance P. The information symbols are then modulated with precoding vectors $\vec{\mathbf{v}}_1(t) \in \mathbb{C}^{m_1(n)}$ at times $t = 1, 2, \ldots, n$. Note that the precoding vector $\vec{\mathbf{v}}_1(t)$ depends only upon the outcome of $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{t-1}$ due to the delayed channel knowledge constraint:

$$\vec{v}_1(t) = f_{\text{signal},1,t}^{(n)}\left(\mathcal{G}_0^{t-1}\right). \tag{3.22}$$

In addition, Tx_1 is allowed to use a vector $\vec{\mathbf{w}}_1 \in \mathbb{C}^{m_2(n)}$ of $m_2(n) \in \mathbb{N}$ noise symbols , which are not necessarily to the interest of any receiver, but can help drown $\vec{\mathbf{x}}$ in the received signal of $\text{Rx}_2, \ldots, \text{Rx}_{k+1}$ such that they cannot decode the message. Each of the noise symbols is a Gaussian random variable with variance P. The noise symbols are also modulated with precoding vectors $\vec{\mathbf{u}}_1(t) \in \mathbb{C}^{m_2(n)}$ at times $t = 1, 2, \ldots, n$. Note that the precoding vector $\vec{\mathbf{u}}_1(t)$ depends only upon the outcome of $\boldsymbol{\mathcal{G}}_{\mathbf{0}}^{t-1}$ due to the delayed channel knowledge constraint:

$$\vec{u}_1(t) = f_{\text{noise},1,t}^{(n)}\left(\mathcal{G}_0^{t-1}\right). \tag{3.23}$$

Similarly, the jammer (i.e. Tx_2) is allowed to use a vector $\vec{\mathbf{w}}_2 \in \mathbb{C}^{m_3(n)}$ of $m_3(n) \in \mathbb{N}$ noise symbols, independent of $\vec{\mathbf{w}}_1$, which are modulated at time t with precoding vector $\vec{\mathbf{u}}_2(t) \in \mathbb{C}^{m_3(n)}$, where

$$\vec{u}_2(t) = f^{(n)}_{\text{noise},2,t}\left(\mathcal{G}_0^{t-1}\right). \tag{3.24}$$

Based on this linear precoding, Tx_1 will then send $\mathbf{x}_1(t) = \vec{\mathbf{v}}_1(t)^\top \vec{\mathbf{x}} + \vec{\mathbf{u}}_1(t)^\top \vec{\mathbf{w}}_1$, and Tx_2 will send $\mathbf{x}_2(t) = \vec{\mathbf{u}}_2(t)^\top \vec{\mathbf{w}}_2$ at time t. We denote the precoding functions used by Tx_1 by $f_1^{(n)} = \{f^{(n)}_{\text{signal},1,t}, f^{(n)}_{\text{noise},1,t}\}_{t=1}^n$, and the ones used by Tx_2 by $f_2^{(n)} = \{f^{(n)}_{\text{noise},2,t}\}_{t=1}^n$. In addition, we denote by $\mathbf{V}_1^n \in \mathbb{C}^{n\times m_1(n)}$, $\mathbf{U}_1^n \in \mathbb{C}^{n\times m_2(n)}$, and $\mathbf{U}_2^n \in \mathbb{C}^{n\times m_3(n)}$ the overall precoding matrices such that the t-th row of $\mathbf{V}_1^n$ is $\vec{\mathbf{v}}_1(t)^\top$, the t-th row of $\mathbf{U}_1^n$ is $\vec{\mathbf{u}}_1(t)^\top$, and the t-th row of $\mathbf{U}_2^n$ is $\vec{\mathbf{u}}_2(t)^\top$.

Based on the above setting, the received signal at Rx_j ($j \in \{1, \ldots, k+1\}$) after the n time steps of the communication will be

$$\vec{\mathbf{y}}_j^n = \mathbf{G}_{j1}^n \mathbf{V}_1^n \vec{\mathbf{x}}_1 + \mathbf{G}_{j1}^n \mathbf{U}_1^n \vec{\mathbf{w}}_1 + \mathbf{G}_{j2}^n \mathbf{U}_2^n \vec{\mathbf{w}}_2 + \vec{\mathbf{z}}_j^n, \tag{3.25}$$

where $\mathbf{G}_{ji}^n$ is the $n\times n$ diagonal matrix whose t-th element on the diagonal is $\mathbf{g}_{ji}(t)$. Now, consider decoding $\vec{\mathbf{x}}$ at Rx_j for $j = 1, \ldots, k+1$. The interference subspace at Rx_j will be

$$\mathcal{I}_j = \text{colspan}\left([\mathbf{G}_{j1}^n \mathbf{U}_1^n \quad \mathbf{G}_{j2}^n \mathbf{U}_2^n]\right), \tag{3.26}$$

where colspan(.) of a matrix is the subspace spanned by its columns, and $[A \quad B]$ denotes the horizontal concatenation of two matrices A, B. Let $\mathcal{I}_j^c \subseteq \mathbb{C}^n$ denote the orthogonal subspace of $\mathcal{I}_j$. Then, in the regime of asymptotically high transmit powers (i.e., ignoring the noise), the decodability of information symbols from Tx_1 at Rx_1 corresponds to the constraints that the image of $\text{colspan}(\mathbf{G}_{11}^n \mathbf{V}_1^n)$ on $\mathcal{I}_1^c$ has dimension $m_1(n)$:

$$\begin{aligned} \dim\left(\text{Proj}_{\mathcal{I}_1^c} \text{colspan}\left(\mathbf{G}_{11}^n \mathbf{V}_1^n\right)\right) &= \dim\left(\text{colspan}\left(\mathbf{V}_1^n\right)\right) \\ &= m_1(n), \end{aligned} \tag{3.27}$$

where $\text{Proj}_{\mathcal{I}_1^c}\text{colspan}\left(\mathbf{G}_{11}^n\mathbf{V}_1^n\right)$ is the orthogonal projection of column span of $\mathbf{G}_{11}^n\mathbf{V}_1^n$ on $\mathcal{I}_1^c$.

Based on this setting, we now define the linear secure degrees of freedom (LSDoF) of the blind cooperative wiretap channel with delayed CSIT.

Definition 8. *d secure degrees of freedom are linearly achievable if there exists a sequence $\{f_1^{(n)}, f_2^{(n)}\}_{n=1}^{\infty}$ such that for each n, $\mathbf{V}_1^n$ satisfies the decodability condition of (3.27) with probability 1, and*

$$d = \lim_{n\to\infty}\frac{m_1(n)}{n}, \tag{3.28}$$

and (Equivocation Condition):

$$\lim_{n\to\infty}\frac{\dim\left(\text{Proj}_{\mathcal{I}_j^c}\text{colspan}\left(\mathbf{G}_{j1}^n\mathbf{V}_1^n\right)\right)}{n} \overset{a.s.}{=} 0, \quad 2 \le j \le k+1. \tag{3.29}$$

*We define $\mathcal{D}$ to be the set of all achievable d's. We also define linear secure degrees of freedom (*LSDoF*) to be the supremum of all $d \in \mathcal{D}$.*

Remark 8. *Equivocation condition in (3.29) implies that $\lim_{P\to\infty}\lim_{n\to\infty}\frac{I(W;\vec{\mathbf{y}}_j^n)}{n\log(P)} = 0, \quad 2 \le j \le k+1$, for linear schemes, where W is the secret message and $\vec{\mathbf{y}}_j^n$ is the received signal at Rx_j; this means that the prelog factor of the Equivocation rate to eavesdroppers would asymptotically vanish as $n \to \infty$.*[9]

The following theorem states that $\frac{1}{3}$ is the maximum secure DoF that can be achieved using linear encoding schemes.

Theorem 5. *For the blind cooperative wiretap channel with a distributed jammer and delayed CSIT,*

$$\text{LSDoF} = \frac{1}{3}. \tag{3.30}$$

[9]This condition is weaker than the condition $\lim_{n\to\infty}\frac{I(W;\vec{\mathbf{y}}_j^n)}{n} = 0, \quad 2 \le j \le k+1$, considered in some prior works. However, one can combine our achievable scheme for blind wiretap channel with delayed CSIT with random binning to satisfy the latter condition as well.

In the case that transmitters have no CSIT with respect to the legitimate receiver (Rx_1), the received signal at all the receivers are statistically the same, and therefore, LSDoF is equal to 0. In addition, in the case that transmitters have instantaneous CSIT with respect to the legitimate receiver, one can show that LSDoF is $\frac{1}{2}$.Therefore, Theorem 5 captures the impact of delayed CSIT as well.

Remark 9. *Theorem 5 implies that no matter how many eavesdroppers exist in the network (as long as there is at least one), the linear secure DoF will be the same.*

In the following sections we provide the proof of achievability and converse for Theorem 5, and explain the key ideas behind the proof.

3.3.2 Proof of Achievability

In this section we prove the achievability for Theorem 5, which characterizes the LSDoF of blind cooperative SISO wiretap channel with delayed CSIT.

Our achievable scheme uses artificial noise alignment to achieve $\frac{1}{3}$. In particular, the scheme keeps the dimension of received signals the same in all receivers, but makes sure the dimension of noise at the legitimate receiver is $\frac{2}{3}$ of that at the eavesdroppers. This way, the legitimate receiver can use $\frac{1}{3}$ of its total received signal dimension to decode its desired message, while the message is completely drowned in noise at the eavesdroppers; because noise at the eavesdroppers will occupy the whole received signal dimension.

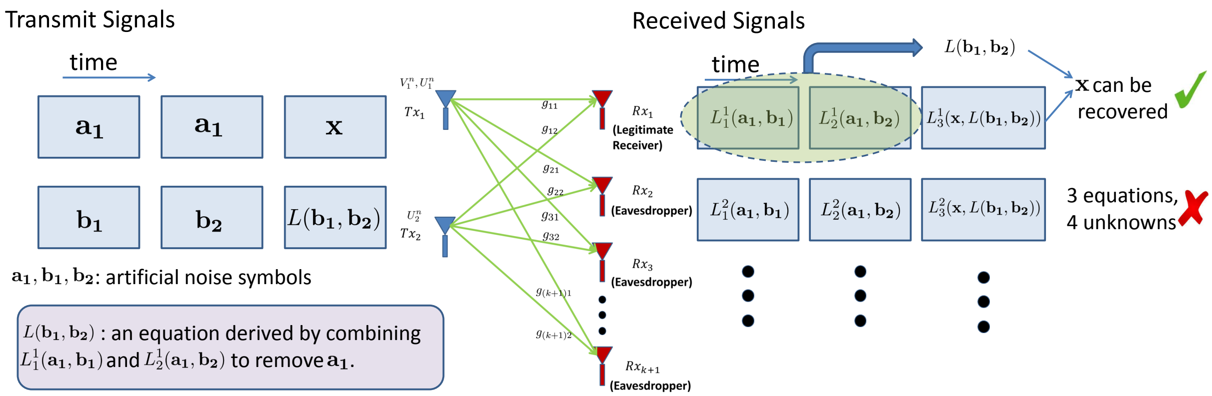


Figure 3.5: The achievable scheme for the blind cooperative wiretap channel with delayed CSIT uses 3 timeslots, where in the first 2 timeslots only artificial noise is being transmitted. In the third timeslot Tx_1 sends the secret symbol x, while Tx_2 sends a noise equation that Rx_1 has already recovered, but not the eavesdroppers.

We set $n = 3$. Let the symbols of the transmitters be denoted by

$$\vec{\mathbf{x}} = [\mathbf{x}], \quad \vec{\mathbf{w}}_1 = [\mathbf{a_1}], \quad \vec{\mathbf{w}}_2 = \begin{bmatrix} \mathbf{b_1} \\ \mathbf{b_2} \end{bmatrix}. \tag{3.31}$$

Transmit symbols are Gaussian random variables with variance P. In $t = 1$, Tx_1 sends the noise symbol $\mathbf{a_1}$, and Tx_2 sends the noise symbol $\mathbf{b_1}$, which results in receiving the following linear combinations at the receivers:

$$\text{Rx}_j : \quad L_1^j(\mathbf{a_1}, \mathbf{b_1}), \qquad j = 2, \ldots, k+1.$$

In $t = 2$, Tx_1 retransmits noise symbol $\mathbf{a_1}$, and Tx_2 sends noise symbol $\mathbf{b_2}$, resulting in the following received signals:

$$\text{Rx}_j : \quad L_2^j(\mathbf{a_1}, \mathbf{b_2}), \qquad j = 2, \ldots, k+1. \tag{3.32}$$

By the end of timeslot 2, Rx_1 has received two equations regarding $\mathbf{a_1}, \mathbf{b_1}, \mathbf{b_2}$;

therefore, it can linearly combine the two equations to remove $\mathbf{a_1}$ and get a new equation $L(\mathbf{b_1}, \mathbf{b_2})$.

In $t = 3$, Tx_1 sends information symbol $\mathbf{x}$, and Tx_2 sends noise equation $L(\mathbf{b_1}, \mathbf{b_2})$, which is already known by Rx_1, but not known by the eavesdroppers almost surely. Therefore, Rx_1 can decode $\mathbf{x}$, while Rx_j, for $j = 2, \dots, k+1$, almost surely cannot decode $\mathbf{x}$. Therefore, $\frac{1}{3}$ linear secure DoF is achieved.

Remark 10. *Note that the above achievable scheme does not depend on how many eavesdroppers exist in the network, hence it implies that* $\text{LSDoF} \geq \frac{1}{3} (\forall k \geq 1)$.

3.3.3 Proof of Converse

Note that LSDoF is non-increasing in the number of eavesdroppers k. Therefore, it is sufficient to show that for the special case of $k = 1$, $\text{LSDoF} \leq \frac{1}{3}$.

There are two key ingredients in proving the converse. The first one is the Rank Ratio Inequality (Lemma 1), which can be used to capture how much the minimum ratio of $\text{rank}\begin{bmatrix}\mathbf{G}_{11}^n\mathbf{U}_1^n & \mathbf{G}_{12}^n\mathbf{U}_2^n\end{bmatrix}$ to $\text{rank}\begin{bmatrix}\mathbf{G}_{21}^n\mathbf{U}_1^n & \mathbf{G}_{22}^n\mathbf{U}_2^n\end{bmatrix}$ is. More specifically, by Rank Ratio Inequality, for any linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{U}_1^n, \mathbf{U}_2^n$ as defined in (3.23)-(3.24),[10]

$$\text{rank}\left[\mathbf{G}_{21}^n\mathbf{U}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n\right] \overset{a.s.}{\leq} \frac{3}{2}\,\text{rank}\left[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n\right]. \tag{3.33}$$

The second ingredient of the converse is the linear version of Least Alignment Lemma, originally presented for general encoding schemes in Lemma 9.

[10]Note that Lemma 1 is stated for the case where transmitters have delayed CSIT of all the channels; however, the same analysis and result holds when transmitters have delayed CSIT with respect to the channels of only Rx_1.

It captures the impact of asymmetric CSIT in the network, and we prove later in Section 3.3.3.

Lemma 12. (Least Alignment Lemma) *For any linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_1^n, \mathbf{U}_1^n, \mathbf{U}_2^n$ *as defined in (3.22)-(3.24),*

$$\text{rank}\left[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n\right] \overset{a.s.}{\leq} \text{rank}\left[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{21}^n\mathbf{U}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n\right].$$

Remark 11. *Lemma 12 implies that when using linear schemes, once the transmitters in a network have no CSIT with respect to a certain receiver, the least amount of alignment will occur at that receiver, meaning that transmit signals will occupy the maximal signal dimensions at that receiver.*

We will now prove the converse using Rank Ratio Inequality and Least Alignment Lemma. First, we state the following claim which can be proved using simple linear algebra, and hence the proof is omitted for brevity.

Claim 1. *For two matrices* A, B *of the same row size,*

- $\text{rank}[A \quad B] - \text{rank}[B] = \dim(\text{Proj}_{\text{colspan}(B)^c}\text{colspan}(A))$;
- $\text{rank}[A \quad B] - \text{rank}[B] = \dim(\text{span}([\vec{s} \quad \vec{0}] \quad | \quad [\vec{s} \quad \vec{0}] \in \text{rowspan}[A \quad B]))$.

Using the second identity in Claim 1, and using some simple linear algebra, one can show the following Corollary.

Corollary 2. *Consider four matrices* A, B, C, D*, where* A, B *have the same number of rows;* C, D *have the same number of rows;* A, C *have the same number of columns; and* B, D *have the same number of columns. Then,*

$$\text{rank}[A \quad B] - \text{rank}[B] \leq \text{rank}[A \quad B; C \quad D] - \text{rank}[B; D], \tag{3.34}$$

where $[B; D]$ *denotes the vertical concatenation of matrices* B *and* D *(i.e.,* $\begin{bmatrix} B \\ D \end{bmatrix}$).

Using Claim 1, the decodability condition in (3.27) can be rewritten as

$$\operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] - \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] = \operatorname{rank}[\mathbf{V}_1^n] = m_1(n). \tag{3.35}$$

Suppose $d \in \mathcal{D}$, i.e., there exists a sequence $\{f_1^{(n)}, f_2^{(n)}\}_{n=1}^{\infty}$ resulting in satisfying (3.27), (3.29) with probability 1, and $d = \lim_{n\to\infty} \frac{m_1(n)}{n}$. Hence, for each n, by the decodability condition in (3.35) we have

$$\operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] - \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] \overset{a.s.}{=} \operatorname{rank}[\mathbf{V}_1^n] \overset{a.s.}{=} m_1(n). \tag{3.36}$$

Furthermore, we define

$$eaves(n) \triangleq \operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{21}^n\mathbf{U}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n] - \operatorname{rank}[\mathbf{G}_{21}^n\mathbf{U}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n]. \tag{3.37}$$

It is easy to see that by Claim 1,

$$eaves(n) = \dim\left(\operatorname{Proj}_{\mathcal{I}_j^c}\operatorname{colspan}\left(\mathbf{G}_{j1}^n\mathbf{V}_1^n\right)\right),$$

with $j = 2$, where I_j is defined in (3.26). Therefore, by Equivocation in (3.29), we have

$$\lim_{n\to\infty} \frac{eaves(n)}{n} \overset{a.s.}{=} 0. \tag{3.38}$$

Therefore, we obtain

$$\begin{aligned} m_1(n) + \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] &\overset{(3.36)}{\underset{a.s.}{=}} \operatorname{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] \\ &\overset{(\text{Lemma 12})}{\underset{a.s.}{\leq}} \operatorname{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{21}^n\mathbf{U}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n] \end{aligned}$$

$$\overset{(3.37)}{\underset{=}{a.s.}} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{U}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n] + eaves(n)$$
$$\overset{(\text{Lemma 1})}{\underset{\leq}{a.s.}} \frac{3}{2}\mathrm{rank}[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] + eaves(n)$$

By rearranging the above inequality, we have

$$2m_1(n) - 2 \times eaves(n) \overset{a.s.}{\leq} \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n]. \tag{3.39}$$

On the other hand, by (3.36),

$$m_1(n) + \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] \overset{(3.36)}{\underset{=}{a.s.}} \mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{11}^n\mathbf{U}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n]] \leq n. \tag{3.40}$$

By summing (3.39) and (3.40), we obtain

$$3m_1(n) - 2 \times eaves(n) \overset{a.s.}{\leq} n.$$

By dividing both sides by n and taking the limit ($n \to \infty$) and using (3.38), we finally get $d \leq \frac{1}{3}$, which completes the proof of the converse. ■

Proof of Lemma 12

Let us fix n, and consider a fixed linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with the corresponding $\mathbf{V}_1^n \in \mathbb{C}^{n\times m_1(n)}, \mathbf{U}_1^n \in \mathbb{C}^{n\times m_2(n)}, \mathbf{U}_2^n \in \mathbb{C}^{n\times m_3(n)}$ as defined in (3.22)-(3.24). For ease of notation, we denote $[\mathbf{V}_1^n \quad \mathbf{U}_1^n]$ by $[\mathbf{W}_1^n]$. Hence, we need to show $\mathrm{rank}[\mathbf{G}_{11}^n\mathbf{W}_1^n \quad \mathbf{G}_{12}^n\mathbf{U}_2^n] \overset{a.s.}{\leq} \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{W}_1^n \quad \mathbf{G}_{22}^n\mathbf{U}_2^n]$. We also define $m \triangleq m_1(n) + m_2(n) + m_3(n)$. We now state a lemma that will be useful later in the proof of Lemma 12.

Lemma 13. **([22])** *A multi-variate polynomial function on $\mathbb{C}^n$ to $\mathbb{C}$, is either identically 0, or non-zero almost everywhere.*

We now prove Lemma 12. Let us denote by $[1 : n]$ the set $\{1, \dots, n\}$. For any matrix $B_{n\times m}$ and $I_1 \subseteq [1 : n]$, and $I_2 \subseteq [1 : m]$, we denote by B_{I_1,I_2} the sub-matrix of B whose rows and columns are specified by I_1 and I_2, respectively. Define the set of realizations $\mathcal{A}$ as:

$$\mathcal{A} \triangleq \{\mathcal{G}^n | \text{rank}[G_{11}^n W_1^n \quad G_{12}^n U_2^n] > \text{rank}[G_{21}^n W_1^n \quad G_{22}^n U_2^n]\}.$$

Note that in order to prove $\text{rank}[\mathbf{G}_{11}^n \mathbf{W}_1^n \quad \mathbf{G}_{12}^n \mathbf{U}_2^n] \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{21}^n \mathbf{W}_1^n \quad \mathbf{G}_{22}^n \mathbf{U}_2^n]$, we only need to show $\Pr(\mathcal{A}) = 0$.

Since a matrix $B_{n\times m}$ has rank r if and only if the maximum size of a square sub-matrix of B with non-zero determinant is r, we have,

$$\begin{aligned}\mathcal{A} \subseteq \{\mathcal{G}^n | \exists I_1 \subseteq [1:n], I_2 \subseteq [1:m], |I_1| = |I_2|, \quad s.t.\\ det([G_{11}^n W_1^n \quad G_{12}^n U_2^n]_{I_1,I_2}) \neq 0, det([G_{21}^n W_1^n \quad G_{22}^n U_2^n]_{I_1,I_2}) = 0\},\end{aligned}$$

which can be rewritten as

$$A \subseteq \cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m]\\ |I_1|=|I_2|}} \{\mathcal{G}^n | det([G_{11}^n W_1^n \quad G_{12}^n U_2^n]_{I_1,I_2}) \neq 0, det([G_{21}^n W_1^n \quad G_{22}^n U_2^n]_{I_1,I_2}) = 0\}. \tag{3.41}$$

Let $X^n \triangleq diag(x_1, \ldots, x_n)$ and $Y^n \triangleq diag(y_1, \ldots, y_n)$, where $x_1, \ldots, x_n, y_1, \ldots, y_n$ are variables in $\mathbb{C}$. Then, for any $I_1 \subseteq [1:n], I_2 \subseteq [1:m]$, where $|I_1| = |I_2|$, $det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2})$ is a multi-variate polynomial function in $x_1, \ldots, x_n, y_1, \ldots, y_n$. Note that if for some realization $X^n = G_{11}^n$ and $Y^n = G_{12}^n$, $det([G_{11}^n W_1^n \quad G_{12}^n U_2^n]_{I_1,I_2}) \neq 0$, then the polynomial function defined by $det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2})$ is not identical to zero ($det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2}) \overset{identical}{\neq} 0$). So, by (3.41), we have

$$\begin{aligned}\mathcal{A} &\subseteq \cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m]\\ |I_1|=|I_2|}} \{\mathcal{G}^n | det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2}) \overset{identical}{\neq} 0, det([G_{21}^n W_1^n \quad G_{22}^n U_2^n]_{I_1,I_2}) = 0\}\\ &= \cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m]\\ |I_1|=|I_2|}} \{\mathcal{G}^n | det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2}) \overset{identical}{\neq} 0,\\ &\qquad G_{21}^n, G_{22}^n \text{ are roots of } det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2})\}.\end{aligned} \tag{3.42}$$

Note that by Lemma 13, for every $I_1 \in [1:n], I_2 \in [1:m], |I_1| = |I_2|$, we have

$$\begin{aligned}
&\Pr(\{\mathcal{G}^n | det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2}) \overset{identical}{\neq} 0, \\
&G_{21}^n, G_{22}{}^n \text{ are roots of } det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2})\}) = 0.
\end{aligned} \tag{3.43}$$

So, since finite union of measure-zero sets has measure zero,

$$\begin{aligned}
&\Pr(\cup_{\substack{I_1 \subseteq [1:n] \\ I_2 \subseteq [1:m] \\ |I_1| = |I_2|}} \{\mathcal{G}^n | det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2}) \overset{identical}{\neq} 0, \\
&G_{21}^n, G_{22}^n\text{: roots of } det([X^n W_1^n \quad Y^n U_2^n]_{I_1,I_2})\}) = 0,
\end{aligned} \tag{3.44}$$

which by (3.42) implies that $\Pr(\mathcal{A}) = 0$. ■

Remark 12. *Using the same line of argument as in the proof of Lemma 12, one can prove Lemma 12 in a more general network setting where there are arbitrary number of transmitters, and the transmitters have arbitrary number of antennas. In addition, the assumption of delayed CSIT of channels to* Rx_1 *can be relaxed to any form of CSIT of channels to* Rx_1 *(e.g. instantaneous CSIT, or partial delayed CSIT). Furthermore, the statement in Lemma 12 holds as long as the number of antennas in* Rx_1 *and* Rx_2 *are equal.*

3.4 Concluding Remarks and Future Directions

In this chapter we focused on information-theoretic secrecy in the context of wiretap channel, where a transmitter wishes to communicate a confidential message to a legitimate receiver in the presence of eavesdroppers. We studied fundamental limits of secure communications when channels are time varying, no CSIT is available with respect to eavesdroppers, and only delayed CSIT is supplied by the legitimate receiver.

We first considered the setting where all nodes in the network are equipped with arbitrary number of antennas, hence called blind MIMOME wiretap channel with delayed CSIT. We completely characterized the secure Degrees of Freedom (SDoF) for all antenna configurations. We strictly improved the existing achievable scheme by proposing a two-phase scheme that utilizes artificial noise alignment to drown the information symbols in noise at the eavesdroppers, while the legitimate receiver can decode all the information symbols. The converse proof is based on four key inequalities used for lower bounding the received signal dimension at receivers which supply different types of CSIT.

We then considered the setting where the secure communication is aided via a distributed jammer, where a jammer is a transmitter that does not necessarily have access to the confidential message, but can help jam the confidential message at the eavesdropper(s). All nodes in the network have a single antenna; hence, this setting was called blind cooperative SISO wiretap channel with delayed CSIT. We characterized the linear SDoF by utilizing the Rank Ratio Inequality (Lemma 1) along with Least Alignment Lemma (Lemma 12), which implies that once the transmitters in a network have no CSIT with respect to a receiver, the least amount of alignment will occur at that receiver, meaning that transmit signals will occupy the maximal signal dimensions at that receiver.

We conjecture that our result on blind cooperative SISO wiretap channel with delayed CSIT is true for general encoding schemes as well. More specifically, if one can prove Conjecture 1 which is stated in Section 2.6, then, together with the generalization of Least Alignment Lemma (Lemma 9), the proof of converse will follow for the general encoding schemes. Another interesting future direction is to study the impact of cooperative jamming on the achievable

SDoF of blind MIMOME wiretap channel with delayed CSIT. Finally, another potential follow-up direction to our work is to consider noisy and delayed CSIT supplied by the legitimate receiver, rather than perfect and delayed CSIT, and study how this impacts the achievable SDoF.

CHAPTER 4

MISO BROADCAST CHANNEL WITH HETEROGENEOUS CSIT

4.1 Overview

In this chapter we continue studying the impacts of channel state information at the transmitters (CSIT) on communication over wireless networks. In particular, we focus on CSIT heterogeneity, which is an important feature of CSIT in real-world wireless networks.[1] The common procedure for obtaining CSIT is to send training symbols (or pilots) at the transmitters, and then estimate the channels at the receivers and feed the estimates back to the transmitters. As a result of this feedback mechanism, CSIT may not always be perfect and instantaneous. For instance, CSIT may be outdated due to the fast fading nature of the channels or slow feedback mechanism, it can be noisy (imperfect), or not available at all. Therefore, one can expect that in a large network there would be various types of CSIT available at the transmitters with respect to different receivers. This results in communication scenarios with *heterogeneous* or *hybrid* CSIT.

Hence, in this chapter we study the impact of heterogeneity of CSIT on the capacity of broadcast channels with a multiple-antenna transmitter and k single-antenna receivers (MISO BC). In particular, we consider the k-user MISO BC, where the CSIT with respect to each receiver can be either instantaneous/perfect, delayed, or not available; and we study the impact of this heterogeneity of CSIT on the degrees-of-freedom (DoF) of such network.

We first focus on the 3-user MISO BC; and we completely characterize the

[1]The results presented in this chapter have been presented in part in [59, 57, 58].

DoF region for all possible heterogeneous CSIT configurations, assuming linear encoding strategies at the transmitters. The result shows that the state-of-the-art achievable schemes in the literature are indeed sum-DoF optimal, when restricted to linear encoding schemes. To prove the result, we develop a novel bound, called *Interference Decomposition Bound*, which provides a lower bound on the interference dimension at a receiver which supplies delayed CSIT based on the average dimension of constituents of that interference, thereby decomposing the interference into its individual components.

Furthermore, we extend our outer bound on the DoF region to the general k-user MISO BC, and demonstrate that it leads to an approximate characterization of linear sum-DoF to within an additive gap of 0.5 for a broad range of CSIT configurations. Moreover, for the special case where only one receiver supplies delayed CSIT, we completely characterize the linear sum-DoF.

4.2 System Model

Throughout this chapter, we use small letters (e.g. x) for scalars, arrowed letters (e.g. $\vec{x}$) for vectors, capital letters (e.g. X) for matrices, and calligraphic font (e.g. $\mathcal{X}$) for sets. We also use bold letters (e.g. $\mathbf{x}$) for random entities, and non-bold letters for deterministic values (e.g., realizations of random variables).

We consider the Gaussian k-user multiple-input single-output broadcast channel (MISO BC) as depicted in Figure 4.1. It consists of a transmitter with m antennas, and k single-antenna receivers, $\mathrm{Rx}_1, \mathrm{Rx}_2, \ldots, \mathrm{Rx}_k$, where $m \geq k$. The transmitter has a separate message for each of the receivers.

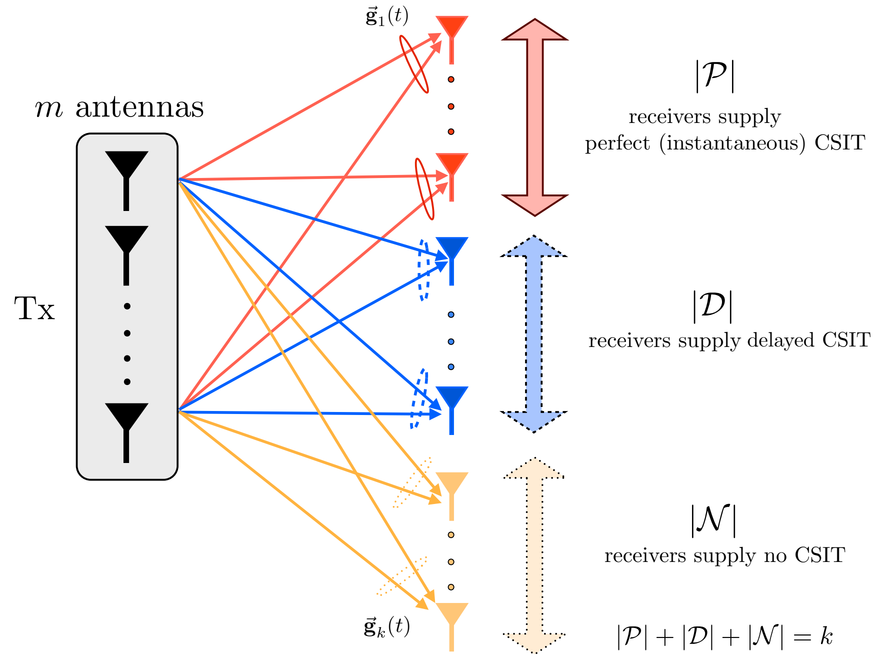


Figure 4.1: Network configuration for k-user MISO BC.

Consider communication over n time slots. The received signal at Rx_j ($j \in \{1, 2, \ldots, k\}$) at time t is given by

$$\mathbf{y}_j(t) = \vec{\mathbf{g}}_j(t)\vec{\mathbf{x}}(t) + \mathbf{z}_j(t), \tag{4.1}$$

where $\vec{\mathbf{x}}(t) \in \mathbb{C}^m$ is the transmit signal vector at time t; $\vec{\mathbf{g}}_j(t) \in \mathbb{C}^{1\times m}$ denotes the channel coefficients of the channel from Tx to Rx_j; and $\mathbf{z}_j(t)$ denotes the additive white noise which is distributed as $\mathcal{CN}(0, 1)$. The elements of the channel coefficients vector $\vec{\mathbf{g}}_j(t)$ are i.i.d, drawn from a continuous distribution and also i.i.d across time and users. $\mathcal{G}(t)$ denotes the set of all k channel vectors at time t. In addition, we denote by $\mathcal{G}^n$ the set of all channel coefficients from time 1 to n, i.e.,

$$\mathcal{G}^n = \{\vec{\mathbf{g}}_j(t) : j = 1, 2, \ldots, k, \quad t = 1, \ldots, n\}. \tag{4.2}$$

We denote the vector of transmit signals in a block of length n by $\vec{\mathbf{x}}^n$, where $\vec{\mathbf{x}}^n$ is the result of concatenation of transmit signal vectors $\vec{\mathbf{x}}(1), \ldots, \vec{\mathbf{x}}(n)$. We assume Tx obeys an average power constraint, $\frac{1}{n}E\{\|\vec{\mathbf{x}}^n\|^2\} \leq P_0$.

We focus on scenarios in which channel state information available at the transmitter (CSIT) with respect to different receivers can be instantaneous (P), delayed (D), or none (N). We refer to these scenarios as *fixed hybrid scenarios*, or *hybrid* in short. In particular, CSIT with respect to Rx_j, $j = 1, 2, \ldots, k$, is denoted by $I_j \in \{P, N, D\}$, as defined in [81]. In this notation, $I_j = P$ indicates that Tx has access to instantaneous CSIT with respect to Rx_j; i.e., at time t, Tx has access to $\{\vec{\mathbf{g}}_j(1), \ldots, \vec{\mathbf{g}}_j(t)\}$. Similarly, $I_j = D$ indicates delayed CSIT with respect to Rx_j; i.e., at time t, Tx has access to $\{\vec{\mathbf{g}}_j(1), \ldots, \vec{\mathbf{g}}_j(t-1)\}$. Finally, $I_j = N$ indicates no CSIT, which means the channel to Rx_j is not known to the Tx at all. We assume that the type of CSIT for each receiver is fixed and does not alternate over time (nevertheless, channels are time-varying). Therefore, there are 3^k different fixed hybrid scenarios. As an example, we use *PDD* to denote the 3-user MISO BC where the first receiver provides instantaneous CSIT, while the other two provide delayed CSIT.

Definition 9. *We denote the set of indices of users in states P, D, N by $\mathcal{P}, \mathcal{D}, \mathcal{N}$, respectively. In addition, for an ordered set $\mathcal{S}$ we denote by $\pi_{\mathcal{S}}$ the ordered set obtained by a permutation of the elements of $\mathcal{S}$, where we denote the elements of the new ordered set by $\pi_{\mathcal{S}}(1), \pi_{\mathcal{S}}(2), \ldots, \pi_{\mathcal{S}}(|\mathcal{S}|)$.*

Note that according to Definition 9, $\mathcal{P} \cup \mathcal{D} \cup \mathcal{N} = \{1, 2, \ldots, k\}$ and $\mathcal{P} \cap \mathcal{D} = \mathcal{D} \cap \mathcal{N} = \mathcal{P} \cap \mathcal{N} = \emptyset$. Based on the above description of channel state information, the channel outcomes available to Tx at time t are denoted by the following set:

$$\tilde{\mathcal{G}}^t = \{\mathbf{G}_i^t; i \in \mathcal{P}\} \cup \{\mathbf{G}_j^{t-1}; j \in \mathcal{D}\}. \tag{4.3}$$

We restrict ourselves to linear coding strategies [52] as defined in Chapter 2, in which degrees-of-freedom (DoF) represents the dimension of the linear subspace of transmitted signals. More specifically, consider a communication scheme with block length n, in which the Tx wishes to deliver a vector $\vec{\mathbf{x}}_j \in \mathbb{C}^{m_j(n)}$ of $m_j(n) \in \mathbb{N}$ information symbols to Rx_j ($j \in \{1, 2. \ldots, k\}$). Each information symbol is a random variable with variance P_0. These information symbols are then modulated with precoding matrices $\mathbf{V}_j(t) \in \mathbb{C}^{m \times m_j(n)}$ at times $t = 1, 2, \ldots, n$. Note that the precoding matrix $\mathbf{V}_j(t)$ depends only upon the outcome of $\tilde{\boldsymbol{\mathcal{G}}}^t$ due to the hybrid CSIT constraint:

$$\mathbf{V}_j(t) = f_{j,t}^{(n)}\left(\tilde{\boldsymbol{\mathcal{G}}}^t\right). \tag{4.4}$$

Based on this linear precoding, Tx will then send $\vec{\mathbf{x}}(t) = \sum_{j=1}^{k} \mathbf{V}_j(t)\vec{\mathbf{x}}_j$ at time t. We can rewrite $\vec{\mathbf{x}}(t)$ as following.

$$\vec{\mathbf{x}}(t) = [\mathbf{V}_1(t) \ldots \mathbf{V}_k(t)][\vec{\mathbf{x}}_1; \ldots; \vec{\mathbf{x}}_k], \tag{4.5}$$

where $[A; B]$ denotes the vertical concatenation of matrices A and B (i.e., $\begin{bmatrix} A \\ \\ B \end{bmatrix}$).

We denote by $\mathbf{V}_j^n \in \mathbb{C}^{nm \times m_j(n)}$ the overall precoding matrix of Tx for Rx_j, such that the rows $1 + (t-1)m, \ldots, tm$ of $\mathbf{V}_j^n$ constitute $\mathbf{V}_j(t)$. In addition, we denote the precoding function used by Tx by $f^{(n)} = \{f_{j,t}^{(n)}\}_{\substack{t=1,\ldots,n \\ j=1,\ldots,k}}$.

Based on the above setting, the received signal at Rx_j ($j \in \{1, 2, \ldots, k\}$) after the n time steps of the communication will be

$$\vec{\mathbf{y}}_j^n = \mathbf{G}_j^n[\mathbf{V}_1^n \ldots \mathbf{V}_k^n][\vec{\mathbf{x}}_1; \ldots; \vec{\mathbf{x}}_k] + \vec{\mathbf{z}}_j^n, \tag{4.6}$$

where $\mathbf{G}_j^n \in \mathbb{C}^{n \times nm}$ is the block diagonal channel coefficients matrix where the channel coefficients of timeslot t (i.e. $\vec{\mathbf{g}}_j(t)$) are in the row t, and in the columns $1 + (t-1)m, \ldots, tm$ of $\mathbf{G}_j^n$, and the rest of the elements of $\mathbf{G}_j^n$ are zero.[2]

Now, consider the decoding of $\vec{\mathbf{x}}_j$ at Rx_j (i.e., decoding the $m_j(n)$ information symbols for Rx_j). The corresponding interference subspace at Rx_j will be

$$\mathcal{I}_j = \mathrm{colspan}\left(\mathbf{G}_j^n[\cup_{i\neq j}\mathbf{V}_i^n]\right),$$

where $[\cup_{i\neq j}\mathbf{V}_i^n]$ is the matrix formed by row concatenation of matrices $\mathbf{V}_i^n$ for $i \neq j$, and colspan(.) of a matrix corresponds to the sub-space that is spanned by its columns. Let $\mathcal{I}_j^{\perp} \subseteq \mathbb{C}^n$ denote the orthogonal subspace of $\mathcal{I}_j$. Then, in the regime of asymptotically high transmit powers (i.e., ignoring the noise), the decodability of information symbols at Rx_j corresponds to the constraint that the image of $\mathrm{colspan}(\mathbf{G}_j^n\mathbf{V}_j^n)$ on $\mathcal{I}_j^{\perp}$ has dimension $m_j(n)$:

$$\dim\left(\mathrm{Proj}_{\mathcal{I}_j^{\perp}}\mathrm{colspan}\left(\mathbf{G}_j^n\mathbf{V}_j^n\right)\right) = \dim\left(\mathrm{colspan}\left(\mathbf{G}_j^n\mathbf{V}_j^n\right)\right) = m_j(n), \tag{4.7}$$

which can be shown by simple linear algebra to be equivalent to the following:

$$\mathrm{rank}[\mathbf{G}_j^n[\cup_{i=1}^k\mathbf{V}_i^n]] - \mathrm{rank}[\mathbf{G}_j^n[\cup_{i\neq j}\mathbf{V}_i^n]] = \mathrm{rank}[\mathbf{G}_j^n\mathbf{V}_j^n] = m_j(n). \tag{4.8}$$

Based on this setting, we now define the linear degrees-of-freedom of the k-user MISO broadcast channel with hybrid CSIT.

Definition 10. *k-tuple $(d_1, d_2, \ldots, d_k)$ degrees-of-freedom are linearly achievable if there exists a sequence $\{f^{(n)}\}_{n=1}^{\infty}$ such that for each n and the corresponding choice of $(m_1(n), m_2(n), \ldots, m_k(n))$, $(\mathbf{V}_1^n, \mathbf{V}_2^n, \ldots, \mathbf{V}_k^n)$ satisfy the decodability condition of (4.8) with probability 1; i.e., for all $j \in \{1, \ldots, k\}$,*

$$\mathrm{rank}[\mathbf{G}_j^n[\cup_{i=1}^k\mathbf{V}_i^n]] - \mathrm{rank}[\mathbf{G}_j^n[\cup_{i\neq j}\mathbf{V}_i^n]] \overset{a.s.}{=} \mathrm{rank}[\mathbf{G}_j^n\mathbf{V}_j^n] \overset{a.s.}{=} m_j(n), \tag{4.9}$$

[2]For $j \in \{1, \ldots, k\}$, we define $\mathbf{G}_j^0[\mathbf{V}_1^0 \ldots \mathbf{V}_k^0] \triangleq \vec{0}$; therefore, for instance, $\mathrm{rank}\left[\mathbf{G}_j^0[\mathbf{V}_1^0 \ldots \mathbf{V}_k^0]\right] = 0$.

and

$$d_j = \lim_{n\to\infty} \frac{m_j(n)}{n}. \tag{4.10}$$

We also define the linear degrees-of-freedom region $\mathrm{LDoF}_{\mathrm{region}}$ *as the closure of the set of all linearly achievable k-tuples* $(d_1, d_2, \ldots, d_k)$. *Furthermore, the linear sum-degrees-of-freedom (*$\mathrm{LDoF}_{\mathrm{sum}}$*) is defined as follows:*

$$\mathrm{LDoF}_{\mathrm{sum}} \triangleq \max \quad \sum_{j=1}^{k} d_j, \qquad s.t. \quad (d_1, d_2, \ldots, d_k) \in \mathrm{LDoF}_{\mathrm{region}}. \tag{4.11}$$

In what follows we first focus on the case of $k = 3$, and completely characterize the $\mathrm{LDoF}_{\mathrm{region}}$ for 3-user MISO BC with hybrid CSIT. We then extend our bounds and present new outer bounds on the $\mathrm{LDoF}_{\mathrm{region}}$ of the general k-user MISO BC with hybrid CSIT.

4.3 3-user MISO Broadcast Channel with Hybrid CSIT

In this section we focus on 3-user MISO broadcast channel with hybrid CSIT. In particular, we first state the complete characterization of $\mathrm{LDoF}_{\mathrm{region}}$ for all hybrid CSIT configurations; and then, we present the proof based on 3 key lemmas.

Theorem 6. *Given a hybrid CSIT configuration, i.e., a partition of* 3 *users into disjoint sets* $\mathcal{P}, \mathcal{D}$, *and* $\mathcal{N}$ *as defined in Definition 9, the* $\mathrm{LDoF}_{\mathrm{region}}$ *is characterized as follows:*

$$\begin{aligned}\mathrm{LDoF}_{\mathrm{region}} = \Big\{ \quad (d_1, d_2, d_3) \quad | \quad & 0 \le d_1, d_2, d_3 \le 1, \\ & \forall i \in \mathcal{D}, \forall \pi_{\mathcal{P}\cup\mathcal{D}\setminus i}, \quad \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{d_{\pi_{\mathcal{P}\cup\mathcal{D}\setminus i}(j)}}{2^j} + d_i + \sum_{j\in\mathcal{N}} d_j \le 1, \\ & \forall \pi_{\mathcal{D}}, \quad \sum_{j\in\mathcal{P}} \frac{d_j}{3} + \sum_{j=1}^{|\mathcal{D}|} \frac{d_{\pi_{\mathcal{D}}(j)}}{j} + \sum_{j\in\mathcal{N}} d_j \le 1,\end{aligned}$$

CSIT States	Linear Degrees of Freedom Region ($\mathrm{LDoF}_{\mathrm{region}}$)	$\mathrm{LDoF}_{\mathrm{sum}}$
PPP	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1\Big\}$	3
PPD	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad \frac{d_1}{2}+\frac{d_2}{4}+d_3 \le 1,\quad \frac{d_1}{4}+\frac{d_2}{2}+d_3 \le 1\Big\}$	$\frac{9}{4}$
PPN	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad d_1+d_3 \le 1,\quad d_2+d_3 \le 1\Big\}$	2
PDD	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad \frac{d_1}{2}+\frac{d_2}{4}+d_3 \le 1,\quad \frac{d_1}{2}+d_2+\frac{d_3}{4} \le 1,\quad \frac{d_1}{3}+\frac{d_2}{2}+d_3 \le 1,\quad \frac{d_1}{3}+d_2+\frac{d_3}{2} \le 1\Big\}$	$\frac{9}{5}$
PDN	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad \frac{d_1}{2}+d_2+d_3 \le 1,\quad d_1+d_3 \le 1\Big\}$	$\frac{3}{2}$
DDD	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad \frac{d_1}{3}+\frac{d_2}{2}+d_3 \le 1,\quad \frac{d_1}{3}+d_2+\frac{d_3}{2} \le 1,\quad \frac{d_1}{2}+\frac{d_2}{3}+d_3 \le 1,\quad \frac{d_1}{2}+d_2+\frac{d_3}{3} \le 1,\quad d_1+\frac{d_2}{2}+\frac{d_3}{3} \le 1,\quad d_1+\frac{d_2}{3}+\frac{d_3}{2} \le 1\Big\}$	$\frac{18}{11}$
DDN	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad \frac{d_1}{2}+d_2+d_3 \le 1,\quad d_1+\frac{d_2}{2}+d_3 \le 1\Big\}$	$\frac{4}{3}$
PNN, DNN, NNN	$\mathrm{LDoF}_{\mathrm{region}} = \Big\{(d_1,d_2,d_3) \mid 0 \le d_1,d_2,d_3 \le 1,\quad d_1+d_2+d_3 \le 1\Big\}$	1

Table 4.1: $\mathrm{LDoF}_{\mathrm{region}}$ and $\mathrm{LDoF}_{\mathrm{sum}}$ for all possible configurations of hybrid CSIT for 3-user MISO BC.

$$\forall i \in \mathcal{P} \cup \mathcal{D}, \quad d_i + \sum_{j \in \mathcal{N}} d_j \le 1 \quad \Big\}. \tag{4.12}$$

The $\mathrm{LDoF}_{\mathrm{region}}$ *and the corresponding* $\mathrm{LDoF}_{\mathrm{sum}}$ *for different CSIT configurations are summarized in Table 4.1.*

Note that although there are 3^3 different CSIT configurations for 3-user MISO BC, many of them are permutations of one another, e.g. PPD, PDP, DPP. As a result, there are only 10 distinct CSIT configurations which are presented in Table 4.1.

Remark 13. *The bound in Theorem 6 strictly improves the state-of-the-art bounds, and also leads to complete characterization of* $\mathrm{LDoF}_{\mathrm{region}}$ *for* $k = 3$*. For instance, for PDD (i.e.* Rx_1 *supplying instantaneous CSIT, while* Rx_2, Rx_3 *supply delayed CSIT) the prior results suggest that* $\mathrm{LDoF}_{\mathrm{sum}} \le \frac{17}{9}$ *[69, 75], while by Theorem 6,* $\mathrm{LDoF}_{\mathrm{sum}}$ *is indeed equal to* $\frac{9}{5}$*. Similarly, for the case of* PPD*, the prior results [69, 75] imply that*

$\mathrm{LDoF}_{\mathrm{sum}} \le \frac{7}{3}$, *while by Theorem 6,* $\mathrm{LDoF}_{\mathrm{sum}} = \frac{9}{4}$

Remark 14. *Theorem 6 implies that the state-of-the-art achievable schemes presented in [9] for PPD and PDD are both optimal from the perspective of* $\mathrm{LDoF}_{\mathrm{sum}}$.

Remark 15. *It is worth noting that in any CSIT configuration which involves receivers with state N, the inequalities that constitute the LDoF region have coefficient 1 for the degrees-of-freedom of receivers with state N. In other words, receivers that supply no CSIT do not contribute to the* $\mathrm{LDoF}_{\mathrm{sum}}$, *and unless all receivers have state N, removing the no CSIT receivers from the network will not decrease the* $\mathrm{LDoF}_{\mathrm{sum}}$.

In the remainder of this section we prove Theorem 6. To this aim, we first present the converse proof in Section 4.3.1, and then discuss the achievability in Section 4.3.2.

4.3.1 Proof of Converse

We first provide the three main ingredients that are key in proving the converse for 3-user MISO broadcast channel with hybrid CSIT. We then show how those main ingredients are used to prove the converse for two representative CSIT configurations (i.e. PDD and PDN). The proof of converse for other CSIT configurations can be found in Appendix C.1. The first two ingredients of the converse proof deal with lower bounding received signal dimension at a receiver which supplies delayed CSIT, while the third ingredient captures the impact of no CSIT.

The first key ingredient is Interference Decomposition Bound, which essentially provides a lower bound on the interference dimension at a receiver supplying delayed CSIT, based on the constituents of that interference, as well as the received signal dimension at other receivers. It is stated below; and its proof is provided in Appendix C.2.

Lemma 14. (Interference Decomposition Bound) *Consider $k = 3$, and a fixed linear coding strategy $f^{(n)}$, with corresponding precoding matrices $\mathbf{V}_1^n, \mathbf{V}_2^n, \mathbf{V}_3^n$ as defined in (4.4). If $I_3 = D$ (i.e., if Rx_3 supplies delayed CSIT),*

$$\frac{\operatorname{rank}[\mathbf{G}_1^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] - \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_2^n] + \operatorname{rank}[\mathbf{G}_3^n\mathbf{V}_2^n]}{2} \overset{a.s.}{\leq} \operatorname{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]]. \tag{4.13}$$

Remark 16. *The R.H.S. of Interference Decomposition Bound represents the dimension of interference caused at Rx_3, which supplies delayed CSIT, by the messages intended for Rx_1, Rx_2. On the other hand, the third term on the L.H.S. (i.e.* $\operatorname{rank}[\mathbf{G}_3^n\mathbf{V}_2^n]$*) is the dimension of the remaining interference at Rx_3 after removing the contribution of the message of Rx_1; and the first two terms (i.e.* $\operatorname{rank}[\mathbf{G}_1^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] - \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_2^n]$*) can be shown by (4.9) and sub-modularity of rank (stated in Lemma 2) to equal* $\operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n]$, *which is the dimension of message of Rx_1. Hence, Interference Decomposition Bound provides an inequality which connects the dimension of interference at a receiver to the average dimension of its constituents. Note that statement of Lemma 14 does not assume any specific CSIT with respect to any receiver except Rx_3.*

The second main ingredient, called MIMO Rank Ratio Inequality for BC, provides a lower bound on the dimension of received signal at a receiver supplying delayed CSIT. It is stated below; and its proof is provided in Appendix C.4.

Lemma 15. (MIMO Rank Ratio Inequality for BC) *Consider $k = 3$, and a linear coding strategy $f^{(n)}$, with corresponding $\mathbf{V}_1^n, \mathbf{V}_2^n, \mathbf{V}_3^n$ as defined in (4.4). If $I_3 = D$ (i.e., if Rx_3 supplies delayed CSIT), then, for each beamforming matrix $\mathbf{V}_i^n$, where $i \in \{1, 2, 3\}$,* *and each $\ell \in \{1, 2, 3\}$, we have*

$$\frac{\text{rank}[[\mathbf{G}_\ell^n; \mathbf{G}_3^n]\mathbf{V}_i^n]}{2} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_3^n \mathbf{V}_i^n], \tag{4.14}$$

where $[\mathbf{G}_\ell^n; \mathbf{G}_3^n]$ denotes the column concatenation of matrices $\mathbf{G}_\ell^n$ and $\mathbf{G}_3^n$.

Remark 17. *Lemma 15 implies that for any transmit signal $\mathbf{V}_i^n$, the corresponding received signal dimension at a receiver with delayed CSIT is at least half of the corresponding received signal dimension at any other receiver. Note that statement of Lemma 15 does not assume any specific CSIT with respect to any receiver except Rx_3.*

The third main ingredient of converse is a variant of Least Alignment Lemma, presented in Lemma 12, for the multi-antenna transmitter configuration. It demonstrates that when using linear schemes, once the transmitter has no CSIT with respect to a certain receiver, the least amount of alignment will occur at that receiver, meaning that transmit signals will occupy the maximal signal dimensions at that receiver. The lemma is stated below. Its proof is similar to the proof of Lemma 12, and is presented in Appendix C.5.

Lemma 16. (Least Alignment Lemma) *Consider $k = 3$, and a linear coding strategy $f^{(n)}$, with corresponding $\mathbf{V}_1^n, \mathbf{V}_2^n, \mathbf{V}_3^n$ as defined in (4.4). For $\mathcal{S} \subseteq \{1, 2, 3\}$ let $\mathbf{V}^n \triangleq [\cup_{i \in \mathcal{S}} \mathbf{V}_i^n]$ denote the row concatenation of the precoding matrices $\mathbf{V}_i^n$, where $i \in \mathcal{S}$. If $I_3 = N$ (i.e., if Rx_3 supplies no CSIT),*

$$\forall \ell \in \{1, 2, 3\}, \qquad \text{rank}\left[\mathbf{G}_\ell^n \mathbf{V}^n\right] \overset{a.s.}{\leq} \text{rank}\left[\mathbf{G}_3^n \mathbf{V}^n\right].$$

Remark 18. *Note that the statement of Lemma 16 does not assume any specific CSIT with respect to any receiver except Rx_3.*

Remark 19. *Lemma 16 is the equivalent of Lemma 9 for the setting where transmitters are restricted to use linear encoding schemes.*

We now prove the converse for two representative CSIT configurations *PDD* and *PDN*, highlighting the applications of the above three lemmas. Converse proofs for other CSIT configurations can be found in Appendix C.1.

Proof of Converse for *PDD*

According to Table 4.1, it is sufficient to show that $\frac{d_1}{2}+\frac{d_2}{4}+d_3 \leq 1$ and $\frac{d_1}{3}+\frac{d_2}{2}+d_3 \leq 1$; since the other two inequalities (i.e. $\frac{d_1}{2}+d_2+\frac{d_3}{4} \leq 1$, and $\frac{d_1}{3}+d_2+\frac{d_3}{2} \leq 1$) can be proven similarly using symmetry. Moreover, the bound $\frac{d_1}{3}+\frac{d_2}{2}+d_3 \leq 1$ follows directly from the existing state-of-the-art arguments used in [69, 65]. Henceforth, we focus on proving $\frac{d_1}{2}+\frac{d_2}{4}+d_3 \leq 1$.

Suppose (d_1, d_2, d_3) degrees-of-freedom are linearly achievable. Hence, by Definition 10 there exists a sequence $\{f^{(n)}\}_{n=1}^{\infty}$ such that for each n and the corresponding choice of $(m_1(n), m_2(n), m_3(n))$, $(\mathbf{V}_1^n, \mathbf{V}_2^n, \mathbf{V}_3^n)$ satisfy the conditions in (4.9) and (4.10). Therefore, in order to prove $\frac{d_1}{2}+\frac{d_2}{4}+d_3 \leq 1$, it is sufficient to show that

$$\frac{m_1(n)}{2}+\frac{m_2(n)}{4}+m_3(n) \overset{a.s.}{\leq} n. \tag{4.15}$$

Note that since in the *PDD* configuration receiver 3 supplies delayed CSIT, we can invoke Lemma 14, which states that:

$$\begin{aligned} 2 \times \text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \;\; \mathbf{V}_2^n]] &\overset{a.s.}{\geq} \text{rank}[\mathbf{G}_1^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] - \text{rank}[\mathbf{G}_1^n\mathbf{V}_2^n] + \text{rank}[\mathbf{G}_3^n\mathbf{V}_2^n] \\ &\overset{(4.9)}{\underset{a.s.}{=}} \text{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + \text{rank}[\mathbf{G}_3^n\mathbf{V}_2^n]. \end{aligned} \tag{4.16}$$

We now further bound each side of the above inequality. We first upper bound the left-hand-side of the above inequality:

$$\text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \overset{(4.9)}{\underset{a.s.}{=}} \text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] - m_3(n) \leq n - m_3(n). \tag{4.17}$$

On the other hand, for the right-hand-side of (4.16) we have

$$
\begin{aligned}
\operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + \operatorname{rank}[\mathbf{G}_3^n\mathbf{V}_2^n] &\overset{\substack{(4.9)\\a.s.}}{=} m_1(n) + \operatorname{rank}[\mathbf{G}_3^n\mathbf{V}_2^n] \\
&\overset{\substack{(\text{Lemma 15})\\a.s.}}{\geq} m_1(n) + \frac{1}{2}\operatorname{rank}[[\mathbf{G}_2^n;\mathbf{G}_3^n]\mathbf{V}_2^n] \\
&\geq m_1(n) + \frac{1}{2}\operatorname{rank}[\mathbf{G}_2^n\mathbf{V}_2^n] \\
&\overset{\substack{(4.9)\\a.s.}}{=} m_1(n) + \frac{1}{2}m_2(n). \qquad (4.18)
\end{aligned}
$$

Hence, by considering (4.16)-(4.18), we obtain

$$
m_1(n) + \frac{1}{2}m_2(n) + 2m_3(n) \overset{a.s.}{\leq} 2n, \qquad (4.19)
$$

which proves (4.15), and therefore, completes the converse proof for *PDD*.

Remark 20. *Note that in order to prove $\frac{d_1}{2} + \frac{d_2}{4} + d_3 \leq 1$ for PDD, we did not rely on any specific CSIT assumption with respect to Rx_2. Therefore, the bound $\frac{d_1}{2} + \frac{d_2}{4} + d_3 \leq 1$ also holds for the case of PPD. Moreover, note that by symmetry one can conclude that $\frac{d_1}{4} + \frac{d_2}{2} + d_3 \leq 1$ also holds for PPD. Hence, since according to Table 4.1, $\frac{d_1}{2} + \frac{d_2}{4} + d_3 \leq 1$ and $\frac{d_1}{4} + \frac{d_2}{2} + d_3 \leq 1$ constitute the LDoF region for PPD, the above derivations suffice in proving the converse for the CSIT configuration PPD as well.*

Proof of Converse for *PDN*

According to Table 4.1, it is sufficient to show that $\frac{d_1}{2} + d_2 + d_3 \leq 1$ and $d_1 + d_3 \leq 1$. Suppose (d_1, d_2, d_3) degrees-of-freedom are linearly achievable. Hence, by Definition 10 there exists a sequence $\{f^{(n)}\}_{n=1}^{\infty}$ such that for each n and the corresponding choice of $(m_1(n), m_2(n), m_3(n))$, $(\mathbf{V}_1^n, \mathbf{V}_2^n, \mathbf{V}_3^n)$ satisfy the conditions in (4.9) and (4.10). Therefore, in order to prove $\frac{d_1}{2} + d_2 + d_3 \leq 1$ and $d_1 + d_3 \leq 1$, it is sufficient to show that

$$
\frac{m_1(n)}{2} + m_2(n) + m_3(n) \overset{a.s.}{\leq} n, \qquad (4.20)
$$

and

$$m_1(n) + m_3(n) \overset{a.s.}{\leq} n. \tag{4.21}$$

We have,

$$\begin{aligned}
\frac{m_1(n)}{2} + m_2(n) + m_3(n) &\overset{(4.9)}{\underset{a.s.}{=}} \frac{\text{rank}[\mathbf{G}_1^n \mathbf{V}_1^n]}{2} + m_2(n) + m_3(n) \\
&\overset{(4.9)}{\underset{a.s.}{=}} \frac{\text{rank}[\mathbf{G}_1^n \mathbf{V}_1^n]}{2} + \text{rank}[\mathbf{G}_2^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \\
&\quad - \text{rank}[\mathbf{G}_2^n [\mathbf{V}_1^n \quad \mathbf{V}_3^n]] + m_3(n) \\
&\overset{\text{(a)}}{\leq} \frac{\text{rank}[\mathbf{G}_1^n \mathbf{V}_1^n]}{2} + \text{rank}[\mathbf{G}_2^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
&\quad - \text{rank}[\mathbf{G}_2^n \mathbf{V}_1^n] + m_3(n) \\
&\leq \frac{\text{rank}[[\mathbf{G}_1^n ; \mathbf{G}_2^n] \mathbf{V}_1^n]}{2} + \text{rank}[\mathbf{G}_2^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
&\quad - \text{rank}[\mathbf{G}_2^n \mathbf{V}_1^n] + m_3(n) \\
&\overset{(b)}{\underset{a.s.}{\leq}} \text{rank}[\mathbf{G}_2^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + m_3(n) \\
&\overset{(4.9)}{\underset{a.s.}{=}} \text{rank}[\mathbf{G}_2^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \\
&\quad - \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
&\overset{\text{(Lemma 16)}}{\underset{a.s.}{\leq}} \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \leq n,
\end{aligned}$$

where (a) follows from the sub-modularity of rank of matrices (see Lemma 2); and (b) follows from Lemma 15 applied to Rx_2 as the receiver which supplies delayed CSIT. Therefore, the proof of (4.20) is complete. We now prove (4.21).

$$\begin{aligned}
m_1(n) + m_3(n) &\overset{(4.9)}{\underset{a.s.}{=}} \text{rank}[\mathbf{G}_1^n \mathbf{V}_1^n] + m_3(n) \\
&\overset{(4.9)}{\underset{a.s.}{=}} \text{rank}[\mathbf{G}_1^n \mathbf{V}_1^n] + \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] - \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
&\overset{\text{(Lemma 2)}}{\leq} \text{rank}[\mathbf{G}_1^n \mathbf{V}_1^n] + \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \mathbf{V}_3^n]] - \text{rank}[\mathbf{G}_3^n \mathbf{V}_1^n] \\
&\overset{\text{(Lemma 16)}}{\underset{a.s.}{\leq}} \text{rank}[\mathbf{G}_3^n [\mathbf{V}_1^n \quad \mathbf{V}_3^n]] \leq n,
\end{aligned}$$

which completes the proof of (4.21).

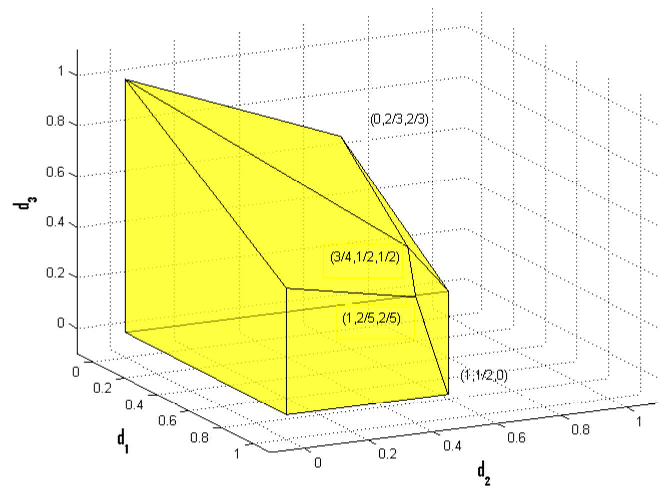


Figure 4.2: LDoF Region for *PDD*.

4.3.2 Proof of Achievability

The regions described in Theorem 6 result in polytopes in $\mathbb{R}^3$; and therefore, the LDoF regions can be completely described via their extreme points. Many of such extreme points can be trivially achieved (e.g. the point $(1, 1, 0)$ for *PPD*); therefore, we only focus on the non-trivial extreme points and provide reference for each of them in Table 4.2.

The only non-trivial extreme point that has not yet been achieved in the literature according to Table 4.2 belongs to *PDD*, and is $(\frac{3}{4}, \frac{1}{2}, \frac{1}{2})$. The LDoF region suggested by Theorem 6 for *PDD* is shown in Fig. 4.2. Therefore, we only prove the achievability of $(\frac{3}{4}, \frac{1}{2}, \frac{1}{2})$ for *PDD*. The scheme is illustrated in Fig. 4.3. We will show how to deliver 3 symbols (a_1, a_2, a_3) to Rx_1, 2 symbols (b_1, b_2) to Rx_2, and 2 symbols (c_1, c_2) to Rx_3 over 4 time slots in order to achieve $(d_1, d_2, d_3) = (\frac{3}{4}, \frac{1}{2}, \frac{1}{2})$.

At $t = 1$, we simply send the uncoded 3 symbols (a_1, a_2, a_3), which are desired by Rx_1. Therefore, the transmit and received signals at the receivers are as

CSIT States	Non-trivial extreme points of the LDoF region and reference to the achievable scheme
PPD	$\left(1, 0, \frac{1}{2}\right), \left(0, 1, \frac{1}{2}\right)$ achieved in Section III-A of [82] $\left(1, 1, \frac{1}{4}\right)$ achieved in Section IV-D of [9]
PDD	$\left(1, 0, \frac{1}{2}\right), \left(0, 1, \frac{1}{2}\right)$ achieved in Section III-A of [82] $\left(1, \frac{2}{5}, \frac{2}{5}\right)$ achieved in Section IV-C of [9] $\left(\frac{3}{4}, \frac{1}{2}, \frac{1}{2}\right)$ achieved in Section 4.3.2 of this chapter $\left(0, \frac{2}{3}, \frac{2}{3}\right)$ achieved in Section III-A of [65]
PDN	$\left(1, \frac{1}{2}, 0\right)$ achieved in Section III-A of [82]
DDD	$\left(\frac{2}{3}, \frac{2}{3}, 0\right), \left(\frac{2}{3}, 0, \frac{2}{3}\right), \left(0, \frac{2}{3}, \frac{2}{3}\right)$ achieved in Section III-A of [65] $\left(\frac{6}{11}, \frac{6}{11}, \frac{6}{11}\right)$ achieved in Section III-B of [65]
DDN	$\left(\frac{2}{3}, \frac{2}{3}, 0\right)$ achieved in Section III-A of [65]

Table 4.2: Achievability results for extreme points of different configurations of hybrid CSIT for 3-user MISO BC

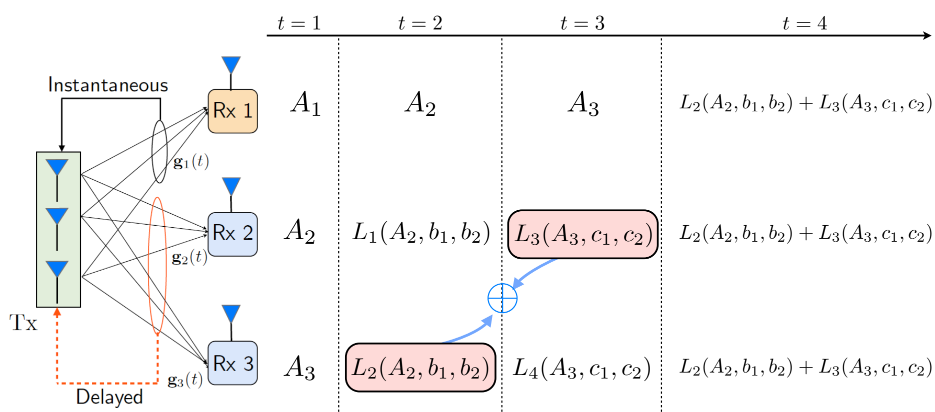


Figure 4.3: Achieving $(d_1, d_2, d_3) = \left(\frac{3}{4}, \frac{1}{2}, \frac{1}{2}\right)$ for *PDD*.

follows (for the sake of DoF analysis, we ignore the additive noise):

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \boldsymbol{y_j}(1) = \vec{\boldsymbol{g}}_j(1) \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad j = 1, 2, 3. \tag{4.22}$$

Denote the linear combinations received by $\mathrm{Rx}_1, \mathrm{Rx}_2, \mathrm{Rx}_3$ at $t = 1$ by A_1, A_2, A_3. Notice that Rx_1 requires A_2, A_3 to be able to (almost surely) decode (a_1, a_2, a_3). Using delayed CSIT from $\mathrm{Rx}_2, \mathrm{Rx}_3$, transmitter can reconstruct A_2, A_3.

At $t = 2$, the transmitter sends the symbols A_2, b_1, b_2 as

$$\vec{\boldsymbol{x}}(2) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} A_2 + \begin{bmatrix} \vec{\boldsymbol{g}}_{\mathbf{1}}(2)^{\perp} \end{bmatrix}^{\top} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \tag{4.23}$$

where $[\vec{\boldsymbol{g}}_{\mathbf{1}}(2)^{\perp}]$ is a 2×3 matrix, where $\vec{\boldsymbol{g}}_{\mathbf{1}}(2)[\vec{\boldsymbol{g}}_{\mathbf{1}}(2)^{\perp}]^{\top} = [0 \quad 0]$. Therefore, the received signals at the Rx_j is $(j = 1, 2, 3)$:

$$\boldsymbol{y}_{\boldsymbol{j}}(2) = \vec{\boldsymbol{g}}_{\boldsymbol{j}}(2) \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} A_2 + \begin{bmatrix} \vec{\boldsymbol{g}}_{\mathbf{1}}(2)^{\perp} \end{bmatrix}^{\top} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right). \tag{4.24}$$

Note that by the end of time slot 2 Rx_1 is able to decode A_2. We denote the linear combinations received by $\mathrm{Rx}_2, \mathrm{Rx}_3$ at $t = 2$ by $L_1(A_2, b_1, b_2), L_2(A_2, b_1, b_2)$, respectively.

At $t = 3$, the transmitted and received signals are:

$$\vec{\boldsymbol{x}}(3) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} A_3 + \begin{bmatrix} \vec{\boldsymbol{g}}_{\mathbf{1}}(3)^{\perp} \end{bmatrix}^{\top} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

$$\boldsymbol{y}_{\boldsymbol{j}}(3) = \vec{\boldsymbol{g}}_{\boldsymbol{j}}(3) \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} A_3 + \begin{bmatrix} \vec{\boldsymbol{g}}_{\mathbf{1}}(3)^{\perp} \end{bmatrix}^{\top} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right),$$

which suggests that Rx_1 would be able to decode A_3. We denote the linear combinations received by $\mathrm{Rx}_2, \mathrm{Rx}_3$ at $t = 3$ by $L_3(A_3, c_1, c_2), L_4(A_3, c_1, c_2)$, respectively.

Note that if Rx_2 is given $L_2(A_2, b_1, b_2)$, it can use its past received signals (i.e., A_2 and $L_1(A_2, b_1, b_2)$) together with $L_2(A_2, b_1, b_2)$ to decode both b_1, b_2 Therefore,

Rx_2 needs $L_2(A_2, b_1, b_2)$. On the other hand, Rx_2 has access to $L_3(A_3, c_1, c_2)$. Similarly, Rx_3 needs $L_3(A_3, c_1, c_2)$ to be able to decode both c_1, c_2, and it has access to $L_2(A_2, b_1, b_2)$. Therefore, at $t = 4$, the transmitter sends $L_2(A_2, b_1, b_2) + L_3(A_3, c_1, c_2)$, which is of interest to both $\mathrm{Rx}_2, \mathrm{Rx}_3$; this is because Rx_3 can then cancel $L_2(A_2, b_1, b_2)$ from its received signal at $t = 4$ to obtain $L_3(A_3, c_1, c_2)$ which it needs. Similarly, Rx_2 can cancel $L_3(A_3, c_1, c_2)$ from its received signal at $t = 4$ to obtain $L_2(A_2, b_1, b_2)$ which it needs. Consequently, all receivers will be able to decode their desired symbols by the end of the fourth time slot; hence, the DoF tuple $(\frac{3}{4}, \frac{1}{2}, \frac{1}{2})$ is achievable. See Fig. 4.3 for an illustration of the achievable scheme.

4.4 k-User MISO BC with Hybrid CSIT

In this section we focus on the general k-user MISO BC with hybrid CSIT. In particular, we first present an outer bound on the LDoF region of the general k-user MISO BC for any arbitrary hybrid CSIT configuration. Then, we show that the bound provides an approximate characterization of $\mathrm{LDoF}_{\mathrm{sum}}$ for the case of $|\mathcal{P}| \geq |\mathcal{D}|$, and exact characterization of $\mathrm{LDoF}_{\mathrm{sum}}$ for $|\mathcal{D}| = 1$. We then present the key tools needed for proving the general outer bound; and finally, we prove the outer bound on the LDoF region.

Theorem 7. *Given a hybrid CSIT configuration, i.e., a partition of k users into disjoint sets $\mathcal{P}, \mathcal{D}$, and $\mathcal{N}$ as defined in Definition 9, the* $\mathrm{LDoF}_{\mathrm{region}}$ *is contained in the following region:*

$$\mathrm{LDoF}_{\mathrm{region}} \subseteq \Big\{ \quad (d_1, \ldots, d_k) \quad | \quad 0 \leq d_1, \ldots, d_k \leq 1,$$

$$\forall i \in \mathcal{D}, \forall \pi_{\mathcal{P}\cup\mathcal{D}\setminus i}, \quad \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{d_{\pi_{\mathcal{P}\cup\mathcal{D}\setminus i}(j)}}{2^j} + d_i + \sum_{j\in\mathcal{N}} d_j \leq 1, \tag{4.25}$$

$$\forall \pi_{\mathcal{D}}, \quad \sum_{j\in\mathcal{P}} \frac{d_j}{k} + \sum_{j=1}^{|\mathcal{D}|} \frac{d_{\pi_{\mathcal{D}}(j)}}{j} + \sum_{j\in\mathcal{N}} d_j \leq 1, \tag{4.26}$$

$$\left.\forall i \in \mathcal{P}\cup\mathcal{D}, \quad d_i + \sum_{j\in\mathcal{N}} d_j \leq 1 \quad \right\}. \tag{4.27}$$

Theorem 7 enables us to approximately characterize LDoF_{sum} to within $\frac{|\mathcal{P}|}{2^{|\mathcal{P}|}}$ for a broad range of CSIT configurations ($|\mathcal{P}| \geq |\mathcal{D}|$). This gap (i.e. $\frac{|\mathcal{P}|}{2^{|\mathcal{P}|}}$) is less than or equal to 0.5, and decays exponentially to zero as $|\mathcal{P}|$ increases. Moreover, Theorem 7 allows us to exactly characterize LDoF_{sum} for the case of $|\mathcal{D}| = 1$. These results are stated more precisely in the following two Propositions.

Proposition 6. *For general k-user MISO BC with $|\mathcal{P}| \geq |\mathcal{D}|$,*

$$|\mathcal{P}| \leq \text{LDoF}_{\text{sum}} \leq |\mathcal{P}| + \frac{|\mathcal{P}|}{2^{|\mathcal{P}|}} \leq |\mathcal{P}| + \frac{1}{2}.$$

Proposition 7. *For general k-user MISO BC with $|\mathcal{D}| = 1$,*

$$\text{LDoF}_{\text{sum}} = |\mathcal{P}| + \frac{1}{2^{|\mathcal{P}|}}. \tag{4.28}$$

Proofs of Propositions 6, 7 are provided in Appendix C.6 and Appendix C.7, respectively. We will now prove Theorem 7. In particular, we first present the key ingredients of the proof, which are the generalizations of Lemmas 14-16. We then prove (4.25)-(4.27).

4.4.1 Key Ingredients for Proof of Theorem 7

Similar to the proof for the case of 3-user MISO BC with hybrid CSIT, we need to extend the Lemmas 14-16. We present the generalizations here, and then prove

Theorem 7. We first present the generalized version of Interference Decomposition Bound in Lemma 14. The proof is provided in Appendix C.2.

Lemma 17. **(Interference Decomposition Bound)** *Consider a fixed linear coding strategy $f^{(n)}$, with corresponding precoding matrices $\mathbf{V}_1^n, \mathbf{V}_2^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4). For any $\mathcal{S} \subseteq \{1, 2, \ldots, k\}$, any $\ell \in \mathcal{S}$, and any $j \notin \mathcal{S}$ for which $I_j = D$,*

$$\frac{\text{rank}[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]] - \text{rank}[\mathbf{G}_\ell^n[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i^n]] + \text{rank}[\mathbf{G}_j^n[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i^n]]}{2} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]], \tag{4.29}$$

where $[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]$ denotes the row concatenation of the corresponding precoding matrices $\mathbf{V}_i^n$, where $i \in \mathcal{S}$.

Remark 21. *Lemma 14 is a special case of Lemma 17 where $\mathcal{S} = \{1, 2\}$, $j = 3$, $\ell = 1$.*

We now present the generalized version of Lemma 15, which is the second main ingredient of the proof, and is proved in Appendix C.4.

Lemma 18. **(MIMO Rank Ratio Inequality for BC)** *Consider a linear coding strategy $f^{(n)}$, with corresponding $\mathbf{V}_1^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4). Let $\mathbf{Y}_j^n \triangleq \mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]$, where $\mathcal{S} \subseteq \{1, 2, \ldots, k\}$. Also, consider distinct receivers $Rx_{i_1}, \ldots, Rx_{i_{j+1}}$, where $j = 1, 2, \ldots, k-1$ and $i_1, \ldots, i_{j+1} \in \{1, \ldots, k\}$. If $Rx_{i_1}, \ldots, Rx_{i_j}$ supply delayed CSIT, then,*

$$\frac{\text{rank}[\mathbf{Y}_{i_1}^n; \ldots; \mathbf{Y}_{i_{j+1}}^n]}{j+1} \overset{a.s.}{\leq} \frac{\text{rank}[\mathbf{Y}_{i_1}^n; \ldots; \mathbf{Y}_{i_j}^n]}{j}. \tag{4.30}$$

Remark 22. *Lemma 15 is a special case of Lemma 18 where $j = 1$, $i_1 = 3, i_2 = \ell$, and $\mathcal{S} = \{i\}$.*

Finally, we present the general version of Lemma 16, which is the third main ingredient for the proof of Theorem 7.

Lemma 19. (Least Alignment Lemma) *For any linear coding strategy $f^{(n)}$, with corresponding $\mathbf{V}_1^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4), and any $\mathcal{S} \subseteq \{1, 2, \ldots, k\}$, if $I_j = N$ for some $j \in \{1, 2, \ldots, k\}$,*

$$\forall \ell \in \{1, 2, \ldots, k\}, \qquad \operatorname{rank}\left[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]\right] \overset{a.s.}{\leq} \operatorname{rank}\left[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]\right],$$

where $[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]$ denotes the row concatenation of the precoding matrices $\mathbf{V}_i^n$, where $i \in \mathcal{S}$.

Using these three ingredients we now proceed to the proof of Theorem 7, and in particular proving the bounds (4.25)-(4.27).

4.4.2 Proof of Bound (4.25) in Theorem 7

Without loss of generality, suppose $\mathcal{P} = \{1, \ldots, |\mathcal{P}|\}$, and $\mathcal{D} = \{|\mathcal{P}|+1, \ldots, |\mathcal{P}|+|\mathcal{D}|\}$, and $\mathcal{N} = \{|\mathcal{P}|+|\mathcal{D}|+1, \ldots, k\}$. In addition, let $i = |\mathcal{P}|+|\mathcal{D}|$, and $\pi_{\mathcal{P}\cup\mathcal{D}\setminus i}$ be the identity permutation. Consequently, we can rewrite (4.25), and our goal is to show

$$\sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{d_j}{2^j} + d_{|\mathcal{P}|+|\mathcal{D}|} + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} d_j \leq 1. \tag{4.31}$$

If the k-tuple $(d_1, d_2, \ldots, d_k)$ degrees-of-freedom are linearly achievable, then by Definition 10 there exists a sequence $\{f^{(n)}\}_{n=1}^{\infty}$ such that for each n and the corresponding choice of $(m_1(n), m_2(n), \ldots, m_k(n))$, $(\mathbf{V}_1^n, \mathbf{V}_2^n, \ldots, \mathbf{V}_k^n)$ satisfy the conditions in (4.9) and (4.10). Therefore, it is sufficient to show

$$\sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{m_j(n)}{2^j} + m_{|\mathcal{P}|+|\mathcal{D}|}(n) + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} n. \tag{4.32}$$

We upper bound each of the three terms on the L.H.S. of (4.32) separately. By induction and application of Lemma 17 and (4.9), one can prove the following claim, which provides an upper bound for the first term on the L.H.S. of (4.32), and is proved in Appendix C.8.

Claim 2.

$$\sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{m_j(n)}{2^j} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]. \tag{4.33}$$

We now upper bound $m_{|\mathcal{P}|+|\mathcal{D}|}(n)$, which is the second term on the L.H.S. of (4.32). By (4.9) we obtain

$$\begin{aligned} m_{|\mathcal{P}|+|\mathcal{D}|}(n) &\overset{a.s.}{=} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\cup_{j=1}^k \mathbf{V}^n_j]] - \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\cup_{j\neq|\mathcal{P}|+|\mathcal{D}|} \mathbf{V}^n_j]] \\ &\overset{\text{(Lemma 2)}}{\leq} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]] - \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]] \\ &\overset{\substack{(a)\\ a.s.}}{\leq} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|+1}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]] - \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]], \end{aligned} \tag{4.34}$$

where (a) follows by Least Alignment Lemma (Lemma 19) since receiver $|\mathcal{P}| + |\mathcal{D}| + 1$ supplies no CSIT.

We now upper bound $\sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^k m_j(n)$, which is the third term on the L.H.S. of (4.32). By (4.9), for all $i \in \{|\mathcal{P}| + |\mathcal{D}| + 1, \dots, k\}$,

$$\begin{aligned} m_i(n) &\overset{a.s.}{=} \text{rank}[\mathbf{G}^n_i[\mathbf{V}^n_1 \dots \mathbf{V}^n_k]] - \text{rank}[\mathbf{G}^n_i[\cup_{j\neq i} \mathbf{V}^n_j]] \\ &\overset{\text{(Lemma 2)}}{\leq} \text{rank}[\mathbf{G}^n_i[\mathbf{V}^n_1 \dots \mathbf{V}^n_i]] - \text{rank}[\mathbf{G}^n_i[\mathbf{V}^n_1 \dots \mathbf{V}^n_{i-1}]]. \end{aligned}$$

Hence, by summing over all the inequalities for $i \in \{|\mathcal{P}|+|\mathcal{D}|+1, \dots, k\}$, we obtain

$$\begin{aligned} \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^k m_j(n) &\overset{a.s.}{\leq} \text{rank}[\mathbf{G}^n_k[\mathbf{V}^n_1 \dots \mathbf{V}^n_k]] - \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|+1}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]] \\ &\quad + \sum_{i=|\mathcal{P}|+|\mathcal{D}|+1}^{k-1} (\text{rank}[\mathbf{G}^n_i[\mathbf{V}^n_1 \dots \mathbf{V}^n_i]] - \text{rank}[\mathbf{G}^n_{i+1}[\mathbf{V}^n_1 \dots \mathbf{V}^n_i]]). \end{aligned} \tag{4.35}$$

Note that since receivers with index in $\{|\mathcal{P}| + |\mathcal{D}| + 1, \dots, k\}$ supply no CSIT, and due to their channel symmetry, for each $i \in \{|\mathcal{P}| + |\mathcal{D}| + 1, \dots, k - 1\}$ we have

$$\text{rank}[\mathbf{G}^n_i[\mathbf{V}^n_1 \dots \mathbf{V}^n_i]] \overset{a.s.}{=} \text{rank}[\mathbf{G}^n_{i+1}[\mathbf{V}^n_1 \dots \mathbf{V}^n_i]]. \tag{4.36}$$

Therefore, by (4.35), (4.36) we obtain

$$\sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_k^n[\mathbf{V}_1^n \dots \mathbf{V}_k^n]] - \text{rank}[\mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|+1}^n[\mathbf{V}_1^n \dots \mathbf{V}_{|\mathcal{P}|+|\mathcal{D}|}^n]]. \tag{4.37}$$

Hence, by summing the inequalities in (4.33), (4.34), and (4.37) we obtain

$$\sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{m_j(n)}{2^j} + m_{|\mathcal{P}|+|\mathcal{D}|}(n) + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_k^n[\mathbf{V}_1^n \dots \mathbf{V}_k^n]] \leq n,$$

which proves (4.32), thus, completing the proof of bound (4.25) in Theorem 7.

4.4.3 Proof of Bound (4.26) in Theorem 7

Without loss of generality, suppose $\mathcal{P} = \{1, \dots, |\mathcal{P}|\}$, and $\mathcal{D} = \{|\mathcal{P}|+1, \dots, |\mathcal{P}|+|\mathcal{D}|\}$, and $\mathcal{N} = \{|\mathcal{P}| + |\mathcal{D}| + 1, \dots, k\}$. In addition, let $\pi_{\mathcal{D}}$ be the reverse of the identity permutation. Consequently, our goal becomes to show

$$\sum_{j=1}^{|\mathcal{P}|} \frac{d_j}{k} + \sum_{j=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|} \frac{d_j}{|\mathcal{P}| + |\mathcal{D}| + 1 - j} + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} d_j \leq 1. \tag{4.38}$$

Suppose $(d_1, \dots, d_k)$ are linearly achievable as defined in Definition 10. Then, by (4.10), it is sufficient to show

$$\sum_{j=1}^{|\mathcal{P}|} \frac{m_j(n)}{k} + \sum_{j=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|} \frac{m_j(n)}{|\mathcal{P}| + |\mathcal{D}| + 1 - j} + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} n. \tag{4.39}$$

We upper bound each of the three terms on the L.H.S. of (4.39) separately. We first upper bound the first term. By (4.9), for all $j = 1, \dots, |\mathcal{P}|$,

$$\begin{aligned}
m_j(n) &\overset{a.s.}{=} \text{rank}[\mathbf{G}_j^n[\mathbf{V}_1^n \dots \mathbf{V}_k^n]] - \text{rank}[\mathbf{G}_j^n[\cup_{i\neq j}\mathbf{V}_i^n]] \\
&\overset{\substack{\text{(Lemma 2)} \\ a.s.}}{\leq} \text{rank}[\mathbf{G}_j^n[\mathbf{V}_1^n \dots \mathbf{V}_j^n]] - \text{rank}[\mathbf{G}_j^n[\mathbf{V}_1^n \dots \mathbf{V}_{j-1}^n]] \\
&\overset{\text{(a)}}{\leq} \text{rank}[[\mathbf{G}_1^n; \dots ; \mathbf{G}_k^n][\mathbf{V}_1^n \dots \mathbf{V}_j^n]] - \text{rank}[[\mathbf{G}_1^n; \dots ; \mathbf{G}_k^n][\mathbf{V}_1^n \dots \mathbf{V}_{j-1}^n]],
\end{aligned}$$

where (a) follows from the fact that for four matrices A, B, C, D, $\text{rank}[A \quad B] - \text{rank}[B] \leq \text{rank}[A \quad B; C \quad D] - \text{rank}[B; D]$, and it can be proven using straightforward linear algebra.

By summing the above inequalities for $j = 1, \ldots, |\mathcal{P}|$, and dividing both sides of the resulting inequality by k we obtain

$$\sum_{j=1}^{|\mathcal{P}|} \frac{m_j(n)}{k} \overset{a.s.}{\leq} \frac{\text{rank}[[\mathbf{G}_1^n; \ldots; \mathbf{G}_k^n][\mathbf{V}_1^n \ldots \mathbf{V}_{|\mathcal{P}|}^n]]}{k}. \tag{4.40}$$

We now upper bound the second term on the L.H.S of (4.39). For the receivers supplying delayed CSIT, i.e. Rx_j, where $j = |\mathcal{P}| + 1, \ldots, |\mathcal{P}| + |\mathcal{D}|$, by (4.9) we have:

$$\begin{aligned} m_j(n) &\overset{a.s.}{=} \text{rank}[\mathbf{G}_j^n[\mathbf{V}_1^n \ldots \mathbf{V}_k^n]] - \text{rank}[\mathbf{G}_j^n[\cup_{i \neq j} \mathbf{V}_i^n]] \\ &\overset{(\text{Lemma } 2)}{\underset{\leq}{a.s.}} \text{rank}[\mathbf{G}_j^n[\mathbf{V}_1^n \ldots \mathbf{V}_j^n]] - \text{rank}[\mathbf{G}_j^n[\mathbf{V}_1^n \ldots \mathbf{V}_{j-1}^n]] \\ &\overset{(b)}{\leq} \text{rank}[[\mathbf{G}_j^n; \ldots; \mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n \ldots \mathbf{V}_j^n]] \\ &\quad -\text{rank}[[\mathbf{G}_j^n; \ldots; \mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n \ldots \mathbf{V}_{j-1}^n]], \end{aligned}$$

where (b) follows from the fact that for four matrices A, B, C, D, $\text{rank}[A \quad B] - \text{rank}[B] \leq \text{rank}[A \quad B; C \quad D] - \text{rank}[B; D]$. Hence, if we divide both sides of the above inequality by $|\mathcal{P}| + |\mathcal{D}| + 1 - j$, and sum over all inequalities for $j = |\mathcal{P}| + 1, \ldots, |\mathcal{P}| + |\mathcal{D}|$, we obtain

$$\begin{aligned} &\sum_{j=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|} \frac{m_j(n)}{|\mathcal{P}| + |\mathcal{D}| + 1 - j} \\ &\overset{a.s.}{\leq} \text{rank}[\mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n[\mathbf{V}_1^n \ldots \mathbf{V}_{|\mathcal{P}|+|\mathcal{D}|}^n] - \frac{\text{rank}[[\mathbf{G}_{|\mathcal{P}|+1}^n; \ldots; \mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n \ldots \mathbf{V}_{|\mathcal{P}|}^n]]}{|\mathcal{D}|} \\ &\quad + \sum_{j=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|-1} \left(\frac{\text{rank}[[\mathbf{G}_j^n; \ldots; \mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n \ldots \mathbf{V}_j^n]]}{|\mathcal{P}| + |\mathcal{D}| + 1 - j} - \frac{\text{rank}[[\mathbf{G}_{j+1}^n; \ldots; \mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n \ldots \mathbf{V}_j^n]]}{|\mathcal{P}| + |\mathcal{D}| - j} \right) \\ &\overset{(\text{Lemma } 18)}{\underset{\leq}{a.s.}} \text{rank}[\mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n[\mathbf{V}_1^n \ldots \mathbf{V}_{|\mathcal{P}|+|\mathcal{D}|}^n] - \frac{\text{rank}[[\mathbf{G}_{|\mathcal{P}|+1}^n; \ldots; \mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n \ldots \mathbf{V}_{|\mathcal{P}|}^n]]}{|\mathcal{D}|} \end{aligned}$$

$$\overset{(\text{Lemma 19})}{\overset{a.s.}{\leq}} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|+1}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}] - \frac{\text{rank}[[\mathbf{G}^n_{|\mathcal{P}|+1}; \dots ; \mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}][\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|}]]}{|\mathcal{D}|}. \tag{4.41}$$

We now upper bound the third term on the L.H.S of (4.39) exactly the same way as we upper bounded the third term on the L.H.S. of (4.32). To avoid redundancy, we only restate the resulting bound which was stated in (4.37).

$$\sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} \text{rank}[\mathbf{G}^n_k[\mathbf{V}^n_1 \dots \mathbf{V}^n_k]] - \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|+1}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]]. \tag{4.42}$$

We now merge the upper bounds on individual terms on the L.H.S. of (4.39). By summing (4.40), (4.41), and (4.42), we obtain

$$\begin{aligned}
&\sum_{j=1}^{|\mathcal{P}|} \frac{m_j(n)}{k} + \sum_{j=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|} \frac{m_j(n)}{|\mathcal{P}| + |\mathcal{D}| + 1 - j} + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \\
\overset{a.s.}{\leq}\ & \text{rank}[\mathbf{G}^n_k[\mathbf{V}^n_1 \dots \mathbf{V}^n_k]] + \frac{\text{rank}[[\mathbf{G}^n_1; \dots ; \mathbf{G}^n_k][\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|}]]}{k} \\
& - \frac{\text{rank}[[\mathbf{G}^n_{|\mathcal{P}|+1}; \dots ; \mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}][\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|}]]}{|\mathcal{D}|} \\
\overset{(c)}{\underset{\leq}{a.s.}}\ & \text{rank}[\mathbf{G}^n_k[\mathbf{V}^n_1 \dots \mathbf{V}^n_k]] \leq n,
\end{aligned} \tag{4.43}$$

where (c) follows from Claim 3, which is stated below and proved in Appendix C.9.

Claim 3.

$$\frac{\text{rank}[[\mathbf{G}^n_1; \dots ; \mathbf{G}^n_k][\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|}]]}{k} \overset{a.s.}{\leq} \frac{\text{rank}[[\mathbf{G}^n_{|\mathcal{P}|+1}; \dots ; \mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}][\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|}]]}{|\mathcal{D}|}. \tag{4.44}$$

Hence, from (4.43), the proof of (4.39) is complete, which concludes the proof of (4.26) in Theorem 7.

4.4.4 Proof of Bound (4.27) in Theorem 7

The proof of (4.27) is similar to proof of (4.25); however, the proof is presented here for completeness. Without loss of generality, suppose $i = |\mathcal{P}| + |\mathcal{D}|$, and $\mathcal{N} = \{|\mathcal{P}| + |\mathcal{D}| + 1, \ldots, k\}$. Consequently, our goal is to show

$$d_{|\mathcal{P}|+|\mathcal{D}|} + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} d_j \leq 1. \tag{4.45}$$

Suppose $(d_1, \ldots, d_k)$ are linearly achievable as defined in Definition 10. Then, by (4.10), it is sufficient to show

$$m_{|\mathcal{P}|+|\mathcal{D}|}(n) + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} n. \tag{4.46}$$

We upper bound each of the two terms on the L.H.S. of (4.46) separately. For the first term on the L.H.S of (4.46), by (4.9), we obtain

$$\begin{aligned} m_{|\mathcal{P}|+|\mathcal{D}|}(n) &\overset{a.s.}{=} \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_k]] - \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\cup_{i \neq |\mathcal{P}|+|\mathcal{D}|} \mathbf{V}^n_i]] \\ &\overset{\substack{\text{(Lemma 2)} \\ a.s.}}{\leq} \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]] - \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]] \\ &\overset{\substack{\text{(Lemma 19)} \\ a.s.}}{\leq} \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|+1}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]] - \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]. \end{aligned} \tag{4.47}$$

We now upper bound the second term on the L.H.S of (4.46), exactly the same way as we upper bounded the third term on the L.H.S. of (4.32). To avoid redundancy, we only restate the resulting bound which was stated in (4.37).

$$\sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} \operatorname{rank}[\mathbf{G}^n_k[\mathbf{V}^n_1 \ldots \mathbf{V}^n_k]] - \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|+1}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|}]]. \tag{4.48}$$

We now sum the upper bounds on individual terms on the L.H.S. of (4.46):

$$m_{|\mathcal{P}|+|\mathcal{D}|}(n) + \sum_{j=|\mathcal{P}|+|\mathcal{D}|+1}^{k} m_j(n) \overset{a.s.}{\leq} \operatorname{rank}[\mathbf{G}^n_k[\mathbf{V}^n_1 \ldots \mathbf{V}^n_k]] - \operatorname{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \ldots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]$$

$$\leq n, \tag{4.49}$$

which completes the proof of (4.46), thus concluding the proof of (4.27) in Theorem 7.

4.5 Concluding Remarks and Future Directions

In this chapter we studied the impact of heterogeneous/hybrid CSIT on the linear DoF (LDoF) of broadcast channels with a multiple-antenna transmitter and k single-antenna receivers (MISO BC), where the CSIT supplied by each receiver can be instantaneous (P), delayed (D), or none (N). We first focused on the 3-user MISO BC; and we completely characterized the DoF region for all possible hybrid CSIT configurations, assuming linear encoding strategies at the transmitter. In order to prove the result, we presented 3 key tools, and in particular, developed a novel bound, called *Interference Decomposition Bound*, which provides a lower bound on the interference dimension at a receiver which supplies delayed CSIT based on the average dimension of constituents of that interference, thereby decomposing the interference into its individual components.

We then extended our main proof ingredients to the general k-user setting; and we presented a general outer bound on linear DoF region of the k-user MISO BC with arbitrary CSIT configuration. We demonstrated how the bound provides an approximate characterization of linear sum-DoF to within an additive gap of 0.5 for the broad range of scenarios in which the number of receivers supplying instantaneous CSIT is greater than the number of receivers supplying delayed CSIT. In addition, for the case where only one receiver supplies delayed CSIT, we completely characterized the linear sum-DoF.

There are several future directions the one can pursue in regards to the work done in this chapter. An interesting direction is to improve both the inner and outer bounds for linear DoF of k-user MISO BC, where the number of receivers supplying instantaneous CSIT is less than the number of receivers supplying delayed CSIT. Another interesting future direction is to extend the results to the non-linear setting (DoF). To this aim, one needs to extend the three main ingredients of the proof of outer bounds to the non-linear setting. The extension of Least Alignment Lemma to the non-linear setting has been presented in Lemma 9, and in [20]. Hence, an interesting direction would be to extend the Interference Decomposition Bound and MIMO Rank Ratio Inequality for BC to the non-linear setting.

CHAPTER 5

TIMELY THROUGHPUT OF HETEROGENEOUS WIRELESS NETWORKS

5.1 Overview

The proliferation of different wireless access technologies, together with the growing number of multi-radio wireless devices suggest that the opportunistic utilization of multiple connections at the users can be an effective solution to the phenomenal growth of traffic demand in wireless networks. This has led to the rise of "heterogeneous" networks. Moreover, since the majority of the traffic volume over wireless networks is time sensitive (e.g. video), timely delivery of messages over wireless networks has become a significant challenge.

In this chapter we study heterogeneous wireless networks with time-sensitive traffic, and provide new insights.[1] In particular, we consider the down-link of a wireless network with N Access Points (AP's) and M clients, where each client is connected to several out-of-band AP's, and requests delay-sensitive traffic (e.g., real-time video). We adopt the framework of Hou, Borkar, and Kumar, and study the maximum total timely throughput of the network, denoted by C_{T^3}, which is the maximum average number of packets delivered successfully before their deadline.

We propose a deterministic relaxation of the problem, which converts the problem to a network with deterministic delays in each link. We show that the additive gap between the capacity of the relaxed problem, denoted by C_{det}, and C_{T^3} is bounded by $2\sqrt{N(C_{\text{det}} + \frac{N}{4})}$, which is asymptotically negligible compared

[1]The results presented in this chapter have been presented in part in [53, 10, 54, 55].

to C_{det}, when the network is operating at high-throughput regime.

In addition, our numerical results show that the actual gap between C_{T^3} and C_{det} is in most cases much less than the worst-case gap proven analytically. Moreover, using LP rounding methods we prove that the relaxed problem can be approximated within additive gap of N. We extend the analytical results to the case of time-varying channel states, real-time traffic, prioritized traffic, and optimal online policies. Finally, we generalize the model for deterministic relaxation to consider fading, rate adaptation, and multiple simultaneous transmissions.

5.2 Network Model and Problem Formulation

In this section we describe our network model and precisely describe the notion of timely throughput introduced in [34]. Finally, we formulate our problem.

5.2.1 Model Setup and Notion of Timely Throughput

We consider the downlink of a network with M wireless clients, denoted by Rx_1, $\mathsf{Rx}_2, \ldots,$ Rx_M, that have packet requests, and N Access Points $\mathsf{AP}_1, \mathsf{AP}_2, \ldots, \mathsf{AP}_N$. These AP's have error-free links to the Backhaul Network (see Fig. 5.1). In addition, time is slotted and transmissions occur during time-slots. Furthermore, the time-slots are grouped into intervals of length τ, where the first interval contains the first τ time-slots, the second interval contains the second τ time-slots, and so on. Moreover, each AP may make one packet transmission in each time-slot.

Each AP is connected via unreliable wireless links to a subset (possibly all) of the wireless clients. These unreliable links are modeled as packet erasure channels that, for now, are assumed to be i.i.d over time, and have fixed success probabilities. In addition, each channel is independent of other channels in the network. (In Section 5.6 these assumptions will be relaxed to consider more general scenarios). The success probability of the channel between AP_i and Rx_j is denoted by p_{ij}, which is the probability of successful delivery of the packet of Rx_j when transmitted by AP_i during a time-slot. If there is no link between an AP and a client, we consider the success probability of the corresponding channel to be 0. Moreover, we assume that the channels do not have interference with each other.

For now we assume that at the beginning of each interval each client has request for a new packet. Right before the start of an interval, each requested packet for that interval is assigned to one of the AP's to be transmitted to its corresponding client. Furthermore, during each time-slot of an interval, each AP picks one of the packets assigned to it to transmit. At the end of that time-slot the AP will know if the packet has been successfully delivered or not. If the packet is successfully delivered, the AP removes that packet from its buffer and does not attempt to transmit it any more. The packets that are not delivered by the end of the interval are dropped from the AP's.

Definition 11. *The decisions on how to assign the requested packets for an interval to the* AP*'s before the start of that interval, and which packet to transmit on a time-slot by each* AP *are specified by a* scheduling policy*. A scheduling policy η makes the decisions causally based on the entire past history of events up to the point of decision-making. We denote the set of all possible scheduling policies by $\mathcal{S}$.*

Definition 12. *A* static scheduling policy, *denoted by η_{static}, is a scheduling policy*

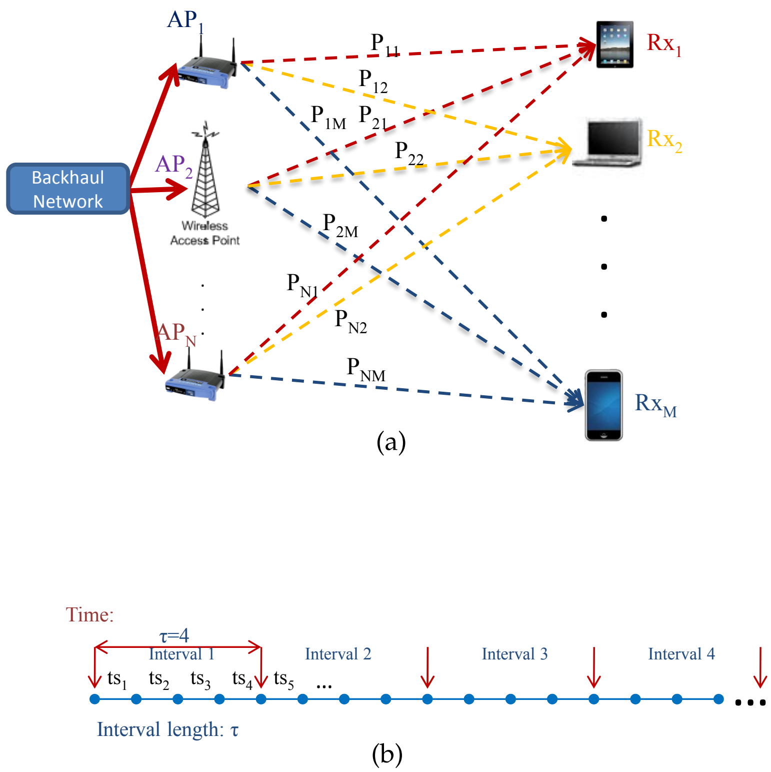


Figure 5.1: Illustration of our network model. Network configuration consisting of N Access points (AP's), M wireless clients, packet erasure channels from AP's to the clients, and the Backhaul network is illustrated in (a). Our time model, in which time is slotted and time-slots are grouped to form intervals of length τ, is shown in (b). In this figure $\tau = 4$.

in which each AP *becomes responsible for serving packets of a fixed subset of clients for all intervals; and the packets of clients assigned to an* AP *are served according to a fixed order. In particular, a static scheduling policy η_{static} is fully specified by a pair $(\vec{\Pi}, \Gamma)$, in which $\vec{\Pi} = [\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_N]$, where $\mathcal{I}_i$'s partition the set $\{1, 2, \ldots, M\}$, indicating how the packet of clients are assigned to* AP*'s. Furthermore, Γ specifies the ordering for the packets assigned to each* AP*. When η_{static} is implemented, each* AP *is responsible for serving packet of the clients assigned to it by $\vec{\Pi}$; and each* AP *persistently transmits a packet until it is delivered successfully, before moving on to the packet of the client with the immediate lower rank in the ordering specified by Γ.*

Definition 13. *A static scheduling policy is called* greedy*, and denoted by $\eta_{g\text{-}static}$, if the order of clients specified by Γ is according to the success probabilities of channels from* AP *to those clients, in decreasing order.*

Assume that a particular scheduling policy η is chosen. For any interval r ($r \in \mathbb{N}$), let $\vec{N}(r, \eta) \triangleq [N_1(r, \eta), N_2(r, \eta), \ldots, N_M(r, \eta)]$ denote the vector of M binary elements whose j^{th} element $N_j(r, \eta)$ is 1 if client Rx_j has successfully received a packet during the r^{th} interval, and 0 otherwise. When using scheduling policy η, the total timely throughput, denoted by $\mathsf{T}^3(\eta)$, is defined as

$$\mathsf{T}^3(\eta) \triangleq \limsup_{r \to \infty} \frac{\sum_{k=1}^{r} \sum_{j=1}^{M} N_j(k, \eta)}{r}. \tag{5.1}$$

In simpler words, $\mathsf{T}^3(\eta)$ is the long-term average number of successful deliveries in the entire network. Similarly, the timely throughput of Rx_j, denoted by $R_j(\eta)$, is defined as

$$R_j(\eta) \triangleq \limsup_{r \to \infty} \frac{\sum_{k=1}^{r} N_j(k, \eta)}{r}, \quad j = 1, 2, \ldots, M. \tag{5.2}$$

Therefore, $R_j(\eta)$ is the long-term average number of successful deliveries for the j^{th} client. Further, we denote the vector of all $R_j(\eta)$'s by $\vec{R}(\eta)$, where we have $\vec{R}(\eta) \triangleq [R_1(\eta), R_2(\eta), \ldots, R_M(\eta)]$. Therefore, the capacity region for timely throughput of M clients in the network is defined as $\mathcal{C} \triangleq \{\vec{R}(\eta) : \eta \in \mathcal{S}\}$.

5.2.2 Main Problem

Our objective is to find the maximum achievable total timely throughput, denoted by C_{T^3}. More precisely, our optimization problem is

$$\textit{Main Problem (MP)}: \qquad C_{\mathsf{T}^3} \triangleq \sup_{\eta \in \mathcal{S}} \mathsf{T}^3(\eta). \tag{5.3}$$

Later in Section 5.6.2 we will consider the problem of finding the maximum weighted total timely throughput $\sum_{j=1}^{M} \omega_j R_j(\eta)$ and its corresponding policy η; but for now we focus on the problem in the case that $\omega_1 = \omega_2 = \ldots = \omega_M = 1$.

5.2.3 Remarks on the Main Problem

As we state later in Lemma 20 in Section 5.4, C_{T^3} can be achieved using a greedy static scheduling policy. Therefore, the optimization in (5.3) can be limited to finding the partition $\vec{\Pi}$ such that the corresponding $\eta_{\text{g-static}}$ maximizes $\mathsf{T}^3(\eta_{\text{g-static}})$. However, this is still quite challenging. In fact, the number of possible greedy static scheduling policies to consider is N^M, which grows exponentially in M.

In [34] Hou et al. have found the timely throughput region for $N = 1$, and have shown that it is a scaled version of a polymatroid [33]. However, when going from one AP to several AP's the problem changes quintessentially: the timely throughput region loses its polymatroidal structure, which makes the problem much more challenging[1]. In this case the timely throughput region is a general polytope with (possibly) exponential number of corner points (corresponding to exponential number of ways of partitioning the clients between the AP's).

[1]Example: Let $N = M = 2, \tau = 1$, and $p_{11} = p_{12} = p_{21} = p_{22} = 1/2$. In this case, the region is the convex hull of three points $(3/4, 0), (1/2, 1/2), (0, 3/4)$. Therefore, no scaled version of the capacity region along its axes can be a polymatroid.

5.3 Deterministic Relaxation and Statement of Main Results

In this section we first explain the intuition behind proposing our relaxation scheme and formulate the relaxed problem. Then, we state the main results.

5.3.1 Deterministic Relaxation

In the system model we assumed channel success probability p_{ij} between AP_i and Rx_j, $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, M$. For now, suppose that $\tau = \infty$, AP_i has only one packet, and wants to transmit that packet to client j. Thus, AP_i persistently sends that packet to client j until the packet goes through. The number of time-slots expended for this packet to be delivered is a Geometric random variable G_{ij} where $\Pr(G_{ij} = k) = p_{ij}(1 - p_{ij})^{k-1}, \quad k \in \mathbb{N}$. We know that $E[G_{ij}] = \frac{1}{p_{ij}}$, and without any deadline, it takes $\frac{1}{p_{ij}}$ time-slots on average for packet of Rx_j to be delivered when transmitted by AP_i.

Therefore, a memory-less erasure channel with success probability p_{ij} can be viewed as a pipe with variable delay which takes a packet from AP_i and gives it to Rx_j according to that variable delay. The probability distribution of the delay is Geometric with parameter p_{ij}.

To simplify the problem, we proposed to relax each channel into a bit pipe with deterministic delay equal to the inverse of its success probability. Therefore, for any packet of Rx_j, when assigned to AP_i for transmission, we associate a fixed size of $\frac{1}{p_{ij}}$ to that packet. This means that each packet assigned to an AP can be viewed as an object with a size, where the size varies from one AP to another; because $\frac{1}{p_{ij}}$'s for different i's are not necessarily the same. On the other hand, we know that each AP has τ time-slots during each interval to send the packets that are assigned to it. Therefore, we can view each AP during each interval as a bin of capacity τ. Therefore, our new problem is a packing problem; i.e., we want to see over all different assignments of objects to bins what the maximum number of objects is that we can fit in those N bins of capacity τ. We denote this maximum possible number of packed objects by C_{det}. More precisely, if we define x_{ij} as the $0-1$ variable which equals 1 if packet of client j is assigned to AP_i, and 0 otherwise, then the relaxed problem can be formulated as following.

Relaxed Problem (RP):

$$C_{\text{det}} \triangleq \max \quad \sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij} \tag{5.4}$$

$$s.t. \quad \sum_{j=1}^{M} \frac{x_{ij}}{p_{ij}} \leq \tau \quad i = 1, 2, \ldots, N \tag{5.5}$$

$$\sum_{i=1}^{N} x_{ij} \leq 1 \quad j = 1, 2, \ldots, M \tag{5.6}$$

$$x_{ij} \in \{0, 1\}. \tag{5.7}$$

5.3.2 Main Results

We now present the main results of the chapter via two Theorems. Theorem 8 bounds the gap between the solution to the main problem (5.3) and its relaxation (5.4). Furthermore, Theorem 9 provides a performance guarantee to the approximation algorithm for the relaxed problem. The proofs of the two Theorems are provided in Section 5.4 and Section 5.5.

Theorem 8. *Let C_{T^3} denote the value of the solution to our main problem in (5.3). Also,* *let C_{det} denote the value of the solution to our relaxed problem in (5.4). We have*

$$C_{det} - 2\sqrt{N(C_{det} + \frac{N}{4})} < C_{\mathsf{T}^3} < C_{det} + N. \tag{5.8}$$

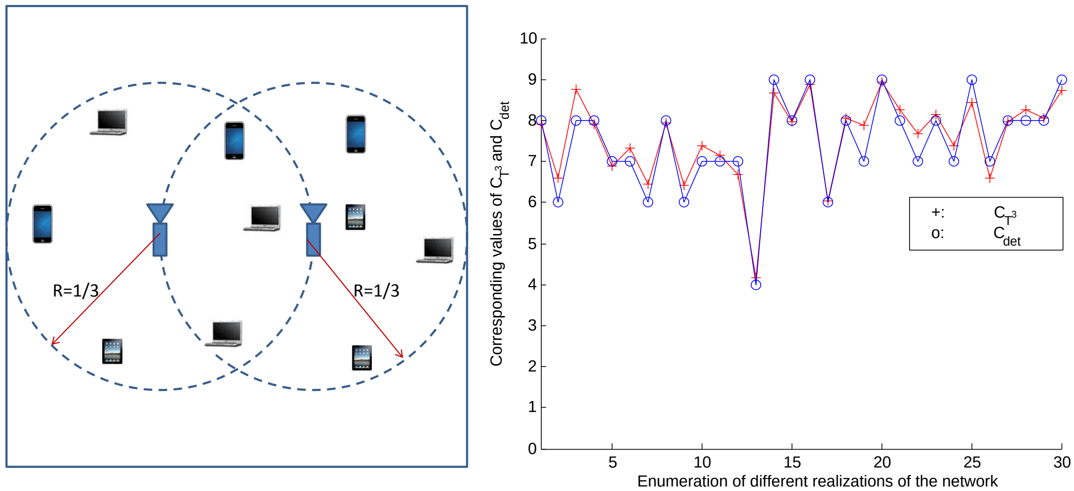
Remark 23. *The right part of the inequality in (5.8) suggests that $C_{\mathsf{T}^3} - C_{det}$ can be no more than N. But the number of* AP*'s N is limited and is usually around $2, 3,$ or 4. Therefore, as $C_{det} \to \infty$ $\frac{N}{C_{det}} \to 0$. Moreover, the left inequality in Theorem 8 suggests that $C_{det} - C_{\mathsf{T}^3}$ becomes negligible compared to C_{det} as $C_{det} \to \infty$. In addition, the inequalities in Theorem 8 imply that as $C_{\mathsf{T}^3} \to \infty$, $C_{det} \to \infty$, too. Therefore, $\frac{C_{det}}{C_{\mathsf{T}^3}} \to 1$, as $C_{\mathsf{T}^3} \to \infty$. Hence, the bounds in Theorem 8 suggest the asymptotic optimality of solving C_{det} instead of C_{T^3}.*

Theorem 8 basically bounds the gap between C_{T^3} and C_{det}. However, a remaining question is: if we run the system based on the greedy static scheduling policy which uses the assignment proposed by the solution to the relaxed problem, how much do we lose in terms of total timely throughput compared to C_{T^3}? The following corollary which is proved in Appendix D.5 addresses this question.

Corollary 3. *Assume $C_{\mathsf{T}^3} \geq \frac{7N}{4}$. Let $\vec{\Pi}_{det}$ denote the assignment of clients to AP's suggested by the solution to the relaxed problem (5.4), and η_{det} be the corresponding greedy static scheduling policy. Then, we have*

$$C_{\mathsf{T}^3} - N - 2\sqrt{N(C_{\mathsf{T}^3} - \frac{3N}{4})} \leq \|\vec{R}(\eta_{det})\|_1 \leq C_{\mathsf{T}^3}.$$

Remark 24. *As we prove in Appendix D.1 the upper bound given in the right inequality in Theorem 8 is tight. Furthermore, the lower bound given in the left inequality of Theorem 1 is tight in terms of order, i.e., there exists a network configuration and a positive constant k for which $C_{det} - C_{\mathsf{T}^3} > k\sqrt{NC_{det}}$.*



(a) (b)

Figure 5.2: Numerical results for the gap between C_{T^3} and C_{det} for the case of two AP's with coverage radius $\frac{1}{3}$, 10 randomly located wireless clients, and intervals of length $\tau = 15$. (a) illustrates the network configuration, where erasure probability of a channel is proportional to the distance between the AP and the corresponding receiver (erasure = $\min\{\frac{\text{distance}}{1/3}, 1\}$). (b) demonstrates the numerical results for the gap for 30 different realizations of the network, where each realization is constructed from a random and uniform location of clients in the network. Each '+' indicates the value of C_{T^3} for each realization, while 'o' indicates the value of C_{det} for the same realization.

Remark 25. *The bounds in Theorem 1 are worst-case bounds, and via numerical experiments we observe that the gap between the original problem and its relaxation is in most cases much smaller. Therefore, the solution to the relaxed problem tracks the solution to the main problem very well, even for a limited number of clients. To illustrate this, consider the network configuration in Figure 5.2(a), where there are two* AP*'s with coverage radius* $\frac{1}{3}$*, and* 10 *clients which are uniformly and randomly located in the coverage area of the two* AP*'s. The erasure probability of the channel between a client and an* AP *is proportional to the distance (erasure* = $\min\{\frac{distance}{1/3}, 1\}$*); and* $\tau = 15$*. For 30 different realizations of this network,* C_{T^3} *and* C_{det} *have been calculated, and plotted in Figure 5.2(b) (detailed numerical results are provided in Section 5.7). The numerical* *results suggest that even for small-scale networks C_{det} is usually very close to C_{T^3}.*

So far, we have shown by Theorem 8 that by considering the relaxed problem (RP) we do not lose much in terms of total timely throughput capacity. Nevertheless, in order for the relaxation to be useful there should be a way to solve the relaxed problem efficiently. The following algorithm approximates the solution to the relaxed problem (RP).

Algorithm 1:

Input: N, M, τ, and p_{ij} for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$.

Find $\mathbf{x}^* = [x^*_{ij}]_{N \times M}$, a basic optimal solution to the LP-relaxation of RP in (5.4).

Output $\lfloor x^*_{ij} \rfloor$ (rounded down version of the elements of $\mathbf{x}^*$) for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$.

The next Theorem, which is proved in Section 5.5, demonstrates that Algorithm 1 approximates the relaxed problem efficiently.

Theorem 9. *Suppose that $\mathbf{x}^*$ is a basic optimal solution to the LP relaxation of RP. We have*

$$C_{det} - \sum_{i=1}^{N} \sum_{j=1}^{M} \lfloor \mathbf{x}^*_{\mathbf{ij}} \rfloor \leq N.$$

Remark 26. *Finding a basic optimal solution to a linear program efficiently is straightforward, and is discussed in [36]. According to Theorem 9 if we find a basic optimal solution to LP relaxation of (5.4), and round down that solution to get integral values, the result will deviate from the optimal solution by at most N. Since N is typically very small (in most cases between 2-4), Algorithm 9 performs well in approximating the solution to the Relaxed Problem (RP).*

Remark 27. *The relaxed problem in (5.4) is a special case of the well-known Maximum Generalized Assignment Problem (GAP). There is a large body of literature on GAP;*

and its special cases capture many combinatorial optimization problems, having several applications in computer science and operations research. Even the special case of GAP in (5.4) is APX-hard [17], meaning that there is no polynomial-time approximation scheme (PTAS) for it.[2] *However, there are several approximation algorithms for GAP, including [17], [44]. In particular, [17], based on a modification of the work in [77], has proposed a 2-approximation algorithm for GAP; and [44] has proposed an LP-based $\frac{e}{e-1}$-approximation algorithm. The performance guarantees in the literature are concerned with multiplicative gap. However, our result in Theorem 9 suggests an additive gap performance guarantee of N for the special case of GAP presented in (5.4). Since N (the number of access points) is typically very small, this provides a tighter approximation guarantee for our problem of interest.*

5.4 Analysis of Approximation Gap (Proof of Theorem 8)

In order to prove Theorem 8, we first state Lemma 20 which is proved in Appendix D.2.

Lemma 20. *C_{T^3} can be achieved using a greedy static scheduling policy.*

Lemma 20 shows there is a scheduling policy which uses the same assignment and ordering of the packets for all intervals, and achieves C_{T^3}. The result in Lemma 20 is intuitive, and is a consequence of time-homogeneity of the system (Lemma 20 is also true for the time-varying channel model where channels are

[2] A PTAS is an algorithm which takes an instance of an optimization problem and a parameter $\epsilon > 0$ and, in polynomial time, produces a solution that is within a factor $1 + \epsilon$ of being optimal (or $1 - \epsilon$ for maximization problems).

modeled by FSMC). In fact, Lemma 20 allows us to focus on only one interval, and then to maximize the expected number of deliveries over that interval.

However, the main challenge lies in how to optimally assign the packets to AP's in order to maximize the expected number of deliveries. But once the assignment is specified, the optimal ordering is trivial according to Lemma 20. We now use Lemma 20 in order to prove the right side of the inequality in Theorem 8.

5.4.1 Proof of $C_{\mathsf{T}^3} < C_{\mathbf{det}} + N$

By Lemma 20 it is sufficient to prove that for any greedy static scheduling policy $\eta_{\text{g-static}}$, $\mathsf{T}^3(\eta_{\text{g-static}}) < C_{\text{det}} + N$. Suppose an arbitrary greedy static scheduling policy $\eta_{\text{g-static}}$ with the corresponding partition $\vec{\Pi}_{\text{g-static}} = [\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_N]$ and ordering $\Gamma_{\text{g-static}}$ is implemented. By (5.1) we know that

$$\mathsf{T}^3(\eta_{\text{g-static}}) = \limsup_{r\to\infty} \frac{\sum_{k=1}^{r}\sum_{j=1}^{M} N_j(k,\eta_{\text{g-static}})}{r}. \tag{5.9}$$

On the other hand, by (5.2) we know that for $j \in [1:M]$, $R_j(\eta_{\text{g-static}}) = \limsup_{r\to\infty} \frac{\sum_{k=1}^{r} N_j(k,\eta_{\text{g-static}})}{r}$. Let Y_i denote the random variable for the number of successful deliveries by AP_i during one interval, when $\eta_{\text{g-static}}$ is implemented; in other words, $Y_i \triangleq \sum_{j\in\mathcal{I}_i} N_j(1,\eta_{\text{g-static}}), \quad i \in [1:N]$. Since a greedy static scheduling policy is implemented and channels are i.i.d over time, by LLN,

$$\begin{aligned}
\sum_{i=1}^{N} E[Y_i] &= \sum_{i=1}^{N} \limsup_{r\to\infty} \frac{\sum_{j\in\mathcal{I}_i}\sum_{k=1}^{r} N_j(k,\eta_{\text{g-static}})}{r} \\
&= \sum_{i=1}^{N}\sum_{j\in\mathcal{I}_i} R_j(\eta_{\text{g-static}}) = \sum_{j=1}^{M} R_j(\eta_{\text{g-static}}) \\
&= \sum_{j=1}^{M} \lim_{r\to\infty} \frac{\sum_{k=1}^{r} N_j(k,\eta_{\text{g-static}})}{r}
\end{aligned}$$

$$= \lim_{r\to\infty} \frac{\sum_{k=1}^{r}\sum_{j=1}^{M} N_j(k, \eta_{\text{g-static}})}{r} = \mathsf{T}^3(\eta_{\text{g-static}}). \tag{5.10}$$

Define $q_i \triangleq |\mathcal{I}_i|$, and denote the enumeration of clients assigned to AP_i by $\{\mathcal{I}_i(1), \mathcal{I}_i(2), \ldots, \mathcal{I}_i(q_i)\}$, where the enumeration is according to the channel success probabilities of different clients in $\mathcal{I}_i$. Let G_{ij} be a geometric random variable with parameter $p_{ij}, i \in [1:N], j \in [1:M]$. Then, it is easy to see that

$$Y_i = \max \quad k \quad s.t. \quad \sum_{j=1}^{k} G_{i\mathcal{I}_i(j)} \le \tau, \quad i \in [1:N], k \le q_i,$$

since $\eta_{\text{g-static}}$ persistently sends a packet until it is delivered, or the interval is over. Define

$$l_i \triangleq \max \quad \hat{l} \quad s.t. \quad \sum_{j=1}^{\hat{l}} 1/p_{i\mathcal{I}_i(j)} \le \tau, \quad \hat{l} \le q_i.$$

Therefore, l_i is the maximum number of objects that fit into bin of capacity τ when the channels are relaxed and clients in $\mathcal{I}_i$ are assigned to AP_i. The following lemma (for which the proof is provided in Appendix D.3) relates l_i to Y_i.

Lemma 21. *Let $\tau \in \mathbb{N}$ and $G_1, G_2, \ldots, G_q$ be independent geometric random variables with parameters $p_1, p_2, \ldots, p_q$ respectively, such that $1 \ge p_1 \ge p_2 \ge \ldots \ge p_q \ge 0$. Also define $l \triangleq \max \hat{l} \quad s.t. \quad \sum_{i=1}^{\hat{l}} 1/p_i \le \tau$, and $Y \triangleq \max i \quad s.t. \quad \sum_{j=1}^{i} G_j \le \tau, \quad i \in \{1, 2, \ldots, q\}$. Then, $E[Y] < l + 1$.*

Hence,

$$\mathsf{T}^3(\eta_{\text{g-static}}) \overset{(a)}{=} \sum_{i=1}^{N} E[Y_i] \overset{(b)}{<} \sum_{i=1}^{N}(l_i + 1) \overset{(c)}{\le} C_{\text{det}} + N.$$

where (a) follows from (5.10); (b) follows from Lemma 21; and (c) follows from the fact that $\sum_{i=1}^{N} l_i$ is the value of the objective function in (5.4) for a feasible solution. Hence the proof of the right inequality in Theorem 8 is complete.

5.4.2 Proof of $C_{\mathbf{det}} - 2\sqrt{N(C_{\mathbf{det}} + \frac{N}{4})} < C_{\mathsf{T}^3}$

Consider the assignment proposed by the solution to the relaxed problem in (5.4), where the clients that are not assigned to any AP for transmission are now assigned to AP's arbitrarily. Let $\vec{\Pi}^{\text{det}}_{\text{g-static}} = [\mathcal{I}^{\text{det}}_1, \mathcal{I}^{\text{det}}_2, \ldots, \mathcal{I}^{\text{det}}_N]$ denote the resulting partition, and also let $\eta^{\text{det}}_{\text{g-static}}$ denote the corresponding greedy static scheduling policy. Therefore, we have $\mathsf{T}^3(\eta^{\text{det}}_{\text{g-static}}) \le C_{\mathsf{T}^3}$. So, it is sufficient to prove that $C_{\text{det}} - 2\sqrt{N(C_{\text{det}} + \frac{N}{4})} < \mathsf{T}^3(\eta^{\text{det}}_{\text{g-static}})$. Let Y^{det}_i denote the random variable indicating the number of successful deliveries by AP_i during one interval, when $\eta^{\text{det}}_{\text{g-static}}$ is implemented, $i = 1, 2, \ldots, N$. With the same argument as in part A we have $\mathsf{T}^3(\eta^{\text{det}}_{\text{g-static}}) = \sum_{i=1}^N E[Y^{\text{det}}_i]$. Therefore, it is sufficient to prove that $C_{\text{det}} - 2\sqrt{N(C_{\text{det}} + \frac{N}{4})} < \sum_{i=1}^N E[Y^{\text{det}}_i]$. Define $q_i = |\mathcal{I}^{\text{det}}_i|$; and denote the enumeration of clients assigned to AP_i by $\{\mathcal{I}^{\text{det}}_i(1), \mathcal{I}^{\text{det}}_i(2), \ldots, \mathcal{I}^{\text{det}}_i(q_i)\}$, where the enumeration is according to the channel success probabilities of different clients in $\mathcal{I}^{\text{det}}_i$. Further, let G_{ij} be a geometric random variable with parameter p_{ij}, $\quad i = 1, 2, \ldots, N, \quad j = 1, 2, \ldots, M$. Then, it is easy to see that

$$Y^{\text{det}}_i = \max \quad k \quad s.t. \sum_{j=1}^{k} G_{i\mathcal{I}^{\text{det}}_i(j)} \le \tau, \quad i \in [1:N], k \le q_i,$$

since $\eta^{\text{det}}_{\text{g-static}}$ persistently sends a packet until it is delivered, or the interval is over. Also define

$$l^{\text{det}}_i \triangleq \max \quad \hat{l} \quad s.t. \quad \sum_{j=1}^{\hat{l}} 1/p_{i\mathcal{I}^{\text{det}}_i(j)} \le \tau, \quad \hat{l} \le q_i.$$

Therefore, l^{det}_i is the maximum number of objects that fit into a bin of capacity τ when the channels are relaxed and clients in $\mathcal{I}^{\text{det}}_i$ are assigned to AP_i. The following lemma (which is proved in Appendix D.4) relates l^{det}_i to Y^{det}_i.

Lemma 22. *Let $\tau \in \mathbb{N}$ and $G_1, G_2, \ldots, G_q$ be independent geometric random variables with parameters $p_1, p_2, \ldots, p_q$ respectively, such that $1 \ge p_1 \ge p_2 \ge \ldots \ge p_q \ge 0$.*

Also define $l \triangleq \max \hat{l} \quad s.t. \quad \sum_{i=1}^{\hat{l}} 1/p_i \le \tau$ *and* $Y \triangleq \max i \quad s.t. \quad \sum_{j=1}^{i} G_j \le \tau, \quad i \in \{1, 2, \ldots, q\}$. *Then,* $l - 2\sqrt{l + \frac{1}{4}} < E[Y]$.

Hence,

$$\begin{aligned}
C_{\mathsf{T}^3} &\ge \sum_{i=1}^{N} E[Y_i^{\text{det}}] \overset{(a)}{>} \sum_{i=1}^{N} l_i^{\text{det}} - 2\sum_{i=1}^{N} \sqrt{l_i^{\text{det}} + \frac{1}{4}} \\
&\overset{(b)}{\ge} \sum_{i=1}^{N} l_i^{\text{det}} - 2\sqrt{N(\sum_{i=1}^{N} l_i^{\text{det}} + \frac{N}{4})} \\
&= C_{\text{det}} - 2\sqrt{N(C_{\text{det}} + \frac{N}{4})},
\end{aligned}$$

where (a) follows from Lemma 22; and (b) follows from Cauchy-Schwarz inequality. Therefore, the left inequality of Theorem 8 is proved and the proof of Theorem 8 is complete.

5.5 Proof of Theorem 9

Note that RP is a mixed integer linear program. Linear relaxation of RP, denoted by LR-RP, replaces the constraint $x_{ij} \in \{0, 1\}$ with $1 \ge x_{ij} \ge 0$ for $i \in [1 : N], \quad j \in [1 : M]$. Any solution to LR-RP can be denoted by an N-by-M matrix $\mathbf{x} = [x_{ij}]_{N \times M}$. So, let $\mathbf{x}^* = [x_{ij}^*]_{N \times M}$ denote a basic optimal solution to LR-RP with objective value V^*; i.e., $V^* = \sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij}^*$. Define

$$\begin{aligned}
Z_1 &= \{j \in \{1, 2, \ldots, M\} | \sum_{i=1}^{N} x_{ij}^* = 0\} \\
Z_2 &= \{j \in \{1, 2, \ldots, M\} | 0 < \sum_{i=1}^{N} x_{ij}^* < 1\} \\
Z_3 &= \{j \in \{1, 2, \ldots, M\} | \sum_{i=1}^{N} x_{ij}^* = 1, \sum_{i=1}^{N} \lfloor x_{ij}^* \rfloor = 0\}
\end{aligned}$$

$$Z_4 = \{j \in \{1, 2, \ldots, M\} | \sum_{i=1}^{N} x^*_{ij} = 1, \sum_{i=1}^{N} \lfloor x^*_{ij} \rfloor = 1\}.$$

It is easy to see that Z_1, Z_2, Z_3, Z_4 partition the set $\{1, 2, \ldots, M\}$. Therefore, $M = |Z_1| + |Z_2| + |Z_3| + |Z_4|$. Furthermore, according to definitions of V^*, Z_1, and Z_4,

$$C_{det} \leq V^* \leq M - |Z_1|, \tag{5.11}$$

$$\sum_{i=1}^{N} \sum_{j=1}^{M} \lfloor x^*_{ij} \rfloor = |Z_4|. \tag{5.12}$$

Hence, by considering (5.11) and (5.12), for proving $C_{det} - \sum_{i=1}^{N} \sum_{j=1}^{M} \lfloor x^*_{ij} \rfloor \leq N$, it is sufficient to prove

$$M - |Z_1| - |Z_4| \leq N, \qquad \text{or equivalently,} \qquad |Z_2| + |Z_3| \leq N. \tag{5.13}$$

We use a similar approach to [83], [12] . Note that since $\mathbf{x} = [x_{ij}]_{N \times M}$ is a basic solution to LR-RP, the number of inequalities in (5.5)-(5.7) tightened by $\mathbf{x}$ is at least the total number of variables, MN. So, if we denote the number of non-tight inequalities in (5.5), (5.6), (5.7) by n_1, n_2, n_3,

$$(N - n_1) + (M - n_2) + (MN - n_3) \geq MN$$

$$\Rightarrow \qquad n_1 + n_2 + n_3 \leq M + N. \tag{5.14}$$

On the other hand, according to definition of Z_1, Z_2, Z_3, Z_4, we have

$$n_1 \geq 0 \tag{5.15}$$

$$n_2 \geq |Z_1| + |Z_2| \tag{5.16}$$

$$n_3 \geq |Z_2| + 2|Z_3| + |Z_4|, \tag{5.17}$$

where (5.17) follows by counting the number of $x^*_{ij} > 0$'s with index j in Z_2, Z_3 or Z_4; the number of $x^*_{ij} > 0$ for which $j \in Z_3$ is at least $2|Z_3|$ since there should be at least two positive fractional values that add up to 1. Hence, by (5.14)-(5.17),

$$|Z_1| + 2|Z_2| + 2|Z_3| + |Z_4| \leq M + N$$

$$\Rightarrow \quad |Z_2| + |Z_3| \leq N,$$

which is the desired inequality as stated in (5.13); therefore, the proof is complete. ■

Corollary 4. *Suppose we choose a basic optimal solution to the LP relaxation of (5.4), denoted by* $\mathbf{x}^*$*, and round down the solution to get integral values. Let* $\vec{\Pi}^{apx}_{det}$ *denote the assignment suggested by the resulting integral values; and let* η^{apx}_{det} *denote the corresponding greedy static scheduling policy. For* $C_{\mathsf{T}^3} > \frac{11N}{4}$ *we have*

$$C_{\mathsf{T}^3} - 2N - 2\sqrt{N(C_{\mathsf{T}^3} - \frac{7N}{4})} \leq \|\vec{R}(\eta^{apx}_{det})\|_1 \leq C_{\mathsf{T}^3}.$$

Proof. Let $C^{\text{apx}}_{\text{det}} \triangleq \sum_{i=1}^{N}\sum_{j=1}^{M}\lfloor x^*_{ij}\rfloor$ denote the objective value of the rounded down basic optimal solution to LR-RP. According to Theorem 8 and Theorem 9, $C^{\text{apx}}_{\text{det}} \geq C_{\text{det}} - N \geq C_{\mathsf{T}^3} - 2N$. Therefore, by using the similar argument as in Corollary 3 the proof will be complete. □

5.6 Extensions

In this section we investigate four important extensions to our main problem formulation: time-varying channels and real-time traffic; weighted total timely throughput; lifting the restriction on splitting packets among AP's; and fading channels, AP's accessing multiple clients simultaneously, clients receiving packets from multiple AP's, and rate adaptation.

5.6.1 Time-Varying Channels and Real-Time Traffic

So far, we have assumed that at the beginning of each interval each client has request for exactly one packet. This assumption can be modified by considering a time-varying packet generation pattern, in which for every interval, each client might have request for no packets, or for multiple packets. In addition, the number of packets requested by clients for one interval might depend on the number of packets requested for other intervals. Furthermore, we have so far assumed that channel success probabilities do not change over time. But, this model can be generalized to include time-varying channels with statistical behaviors that are not necessarily independent of one another.

We capture the above two generalizations by considering an irreducible Finite-State Markov Chain (FSMC), in which each state jointly specifies the number of packets requested by each client, as well as the channel states for different channels during an interval. When a new interval begins, the Markov Chain might change its state, and in this case, packets for a new subset of clients are requested, and the channel reliabilities change. Denote the set of all possible states of the FSMC by C. Each state $\lambda \in C$ specifies a pair $(\vec{B}(\lambda), \mathbf{P}(\lambda))$, where $\vec{B}(\lambda) \triangleq [B_1(\lambda), B_2(\lambda), \ldots, B_M(\lambda)]$, such that $B_j(\lambda)$ is the number of the packets requested by client j, and $\mathbf{P}(\lambda)$ is an $N \times M$ matrix that contains channel success probabilities. It is assumed that channel success probabilities remain the same during each interval, and are known to the AP's.

Our objective is again to find C_{T^3}. We use a similar argument to the one in [30] for extensions to time-varying channels and variable-bit-rate traffic. In particular, we decompose the set of intervals into different subsets, where each subset contains the intervals that are in the same state of the FSMC. For those intervals in which the system is at state λ, we convert our problem to an instance of the problem described in Section 5.2. More particularly, for the system described by state λ, we ignore all the clients that do not have packet request. Furthermore, for any $j \in \{1, 2, \ldots, M\}$ where $B_j(\lambda) > 1$ we consider $B_j(\lambda) - 1$ virtual clients, such that the channel between AP_i and each of those virtual clients would have success probability $P_{ij}(\lambda)$. This means that these virtual clients are copies of Rx_j. Consequently, for the intervals for which the system is at state λ the problem becomes the same as described in Section 5.2. With the same argument as in proof of Theorem 8, there exists a fixed assignment $\vec{\Pi}(\lambda)$, which if used together with its corresponding optimal ordering for such intervals, achieves the optimal T^3 for those intervals. We denote this optimal T^3 by $C_{\mathsf{T}^3}(\lambda)$. In addition, let $C_{\text{det}}(\lambda)$ denote the solution to the relaxed problem when the system is at state λ. For any state $\lambda \in \mathcal{C}$, with the same argument as in the proof of Theorem 8, we have $C_{\text{det}}(\lambda) - 2\sqrt{N(C_{\text{det}}(\lambda) + \frac{N}{4})} < C_{\mathsf{T}^3}(\lambda) < C_{\text{det}}(\lambda) + N$. Now, let π_λ denote the steady state probability of λ. Therefore, $C_{\mathsf{T}^3} = \sum_{\lambda \in \mathcal{C}} \pi_\lambda C_{\mathsf{T}^3}(\lambda), C_{\text{det}} = \sum_{\lambda \in \mathcal{C}} \pi_\lambda C_{\text{det}}(\lambda)$. Hence,

$$C_{\text{det}} - 2\sum_{\lambda \in \mathcal{C}} \pi_\lambda \sqrt{N(C_{\text{det}}(\lambda) + \frac{N}{4})} < C_{\mathsf{T}^3} < C_{\text{det}} + N. \tag{5.18}$$

On the other hand, by using Cauchy-Schwarz inequality we have

$$\begin{aligned}
&\sum_{\lambda \in \mathcal{C}} \pi_\lambda \sqrt{N(C_{\text{det}}(\lambda) + \frac{N}{4})} \\
&\leq \sqrt{\sum_{\lambda \in \mathcal{C}} \pi_\lambda} \sqrt{\sum_{\lambda \in \mathcal{C}} N\pi_\lambda (C_{\text{det}}(\lambda) + \frac{N}{4})} = \sqrt{N(C_{\text{det}} + \frac{N}{4})}.
\end{aligned} \tag{5.19}$$

Putting (5.18) and (5.19) together we get $C_{\text{det}} - 2\sqrt{N(C_{\text{det}} + \frac{N}{4})} < C_{\mathsf{T}^3} < C_{\text{det}} + N$, which is the same as the result in Theorem 8.

Theorem 10. *For the network model described in Section 5.2 consider the extension to time-varying channels and real-time traffic, modeled by the FSMC described in Section*

5.6.1, where each state of FSMC captures both the success probability of channels and the number of packets for each client during an interval. We have

$$C_{det} - 2\sqrt{N(C_{det} + \frac{N}{4})} < C_{\mathsf{T}^3} < C_{det} + N.$$

5.6.2 Weighted Total Timely Throughput

In Section 5.2 we considered the same importance for all the flows in the network; and our objective was to maximize T^3. However, it might be the case that in a network some of the flows are more important than the others, and should be prioritized accordingly. In this section the formulation remains the same as the one described in Section 5.2, except the objective function, which rather than maximizing T^3, maximizes a weighted average of timely throughputs. In particular, weighted total timely throughput of the scheduling policy η, w-$\mathsf{T}^3(\eta)$, is defined as

$$w\text{-}\mathsf{T}^3(\eta) \triangleq \sum_{j=1}^{M} \omega_j R_j(\eta), \tag{5.20}$$

where ω_j's are arbitrary weights greater than 1 ($j = 1, 2, \ldots, M$). Our objective is to find

$$C_{w\text{-}\mathrm{T}^3} \triangleq \sup_{\eta \in \mathcal{S}} \quad w\text{-}\mathsf{T}^3(\eta). \tag{5.21}$$

For this extension of the problem we again propose the channel relaxation which results in a new integer program. This integer program is again a GAP. The formulation of the relaxed problem is as follows:

$$C_{w\text{-det}} \triangleq \max \sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij}\omega_j \tag{5.22}$$

$$\text{s.t.} \quad \sum_{j=1}^{M} \frac{x_{ij}}{p_{ij}} \leq \tau$$

$$\sum_{i=1}^{N} x_{ij} \leq 1$$

$$x_{ij} \in \{0, 1\}, \qquad i = [1 : N], \quad j = [1 : M].$$

The following theorem, which is proved in the Appendix D.6, states that the value of the solution to (5.21) is asymptotically the same as the value of the solution to (5.22) as $C_{w\text{-}\mathsf{T}^3} \to \infty$ (or equivalently $C_{w\text{-det}} \to \infty$).

Theorem 11. *Let $C_{w\text{-}\mathsf{T}^3}$ denote the value of the solution to (5.21). Further, let $C_{w\text{-}det}$ denote the value of the solution to (5.22). Then, for $\omega_{max} = \max\{\omega_1, \omega_2, \ldots, \omega_M\}$,*

$$C_{w\text{-}det} - 2\omega_{max}\sqrt{N(C_{w\text{-}det} + \frac{N}{4})} < C_{w\text{-}\mathsf{T}^3} < C_{w\text{-}det} + N\omega_{max}. \tag{5.23}$$

5.6.3 Dynamic Splitting

We assumed in Section 5.2 that the packets are partitioned between AP's at the beginning of each interval, to reduce the overhead for tracking ACKs and NACKs in the network. If packets are available to all AP's for transmission (i.e., no partitioning is done beforehand), in order to maximize the total timely throughput, each AP has to constantly track all ACKs and NACKs of all clients, in order to know whether a packet has already been delivered to its destination. Here we lift the partitioning restriction to understand how much capacity gain can be obtained. We first describe the model, and formulate the problem as a Markov Decision Process (MDP). We then discuss the tractability of solving the MDP, propose a fast greedy heuristic for the MDP, and analyze its computational complexity. Finally, we show the performance of the proposed heuristic via numerical results.

Network Model

We consider the same network configuration, time model, channel model, and packet arrival as in Section 5.2. Nevertheless, the packets requested for each interval are now available to all AP's (i.e. they are not split among the AP's at the beginning of each interval), and a packet might be served by several AP's. The AP's can then dynamically choose what packet to transmit in a coordinated manner at each time-slot. The choice of the packet to be sent by each AP may be based on the channels and the past outcomes of the transmissions. Our objective is to find a scheduling policy which maximizes the total timely throughput of the system, as defined in Section 5.2. We call the optimal scheduling the "Optimal Online Scheduling", since each AP has to decide what the optimal strategy is at each time-slot.

An MDP Formulation

One can argue in a similar way as in Lemma 1 that due to the time-homogeneous structure of the system, the maximal total timely throughput is equal to the maximum achievable expected number of deliveries in one interval. Therefore, the new problem can be formulated as a finite-horizon Markov Decision Process (MDP), as detailed below:

State Space: The state of the system is an $(M+1)$-tuple where the first M components are binary variables $\{Q_j(t)\}_{j=1}^{M}$, and $Q_j(t) = 1$ if Rx_j has not yet received its packet successfully, and $Q_j(t) = 0$ otherwise. The $(M+1)$-th component is the time-slot that the system is currently at, i.e. $Q_{M+1} \in \{1, 2, \ldots, \tau\}$. We denote the state space by $\mathcal{Q}$.

Action Space: For any state $s \in Q$ corresponding to set of clients $U(s)$ not having received their packets yet, the action space is an N-tuple $(i_1, \ldots, i_N)$ where $i_k \in U(s) \cup \{0\}$ for $k \in \{1, 2, \ldots, N\}$. If $i_k = j$, it means that client j is served by AP_k, and if $j = 0$, AP_k will be idle during the time-slot. A policy $\mathcal{P}$ is a function mapping the state space to action space.

Reward: For successful delivery of each packet, a reward equal to 1 is obtained.

Transition Function: For $t < \tau$, transition probability from state $s = (q_1, \ldots, q_M, t)$ to state
$s' = (q'_1, \ldots, q'_M, t+1)$ using action $a(s)$ is simply probability of the event in which in one time-slot using action $a(s)$ the state changes from s to s'. [1]

Objective: We want to find the optimal policy that maximizes the expected number of deliveries in τ time-slots. The objective is similar to the objective initially considered in Section II.

One can use the common technique of using Dynamic Programming to calculate the maximal value. More specifically, for 2 AP's ($N = 2$), the optimization problem reduces to the following.

Let $V^t(U)$ denote the maximum expected number of deliveries for the set of packets U and during time-slots $t, t+1, \ldots, \tau$. Therefore, the objective can be rewritten as follows.

$$\text{Objective:} \qquad V^1(\{1, 2, \ldots, M\}),$$

where

[1] More specifically, the transition probability is $Pr(\cap_{j=1}^{M}(\cup_{1 \le k \le N, i_k = j} B(p_{kj}) = q_j - q'_j))$, where $B(p)$ is a Bernoulli random variable with parameter p.

$$
\begin{aligned}
V^t(U) = \max_{\{i,j\}\in U} \{ & p_{1i}p_{2j}[2 - \mathbb{I}(i = j) + V^{t+1}(U \setminus \{i, j\})] \\
& + p_{1i}(1 - p_{2j})[1 + V^{t+1}(U \setminus \{i\})] \\
& + (1 - p_{1i})p_{2j}[1 + V^{t+1}(U \setminus \{j\})] \\
& + (1 - p_{1i})(1 - p_{2j})V^{t+1}(U)\},
\end{aligned}
$$

and $V^\tau(U) = \max_{\{i,j\}\in U}[p_{1i} + p_{2j} - p_{1i}p_{2j}\mathbb{I}(i = j)]$, where $\mathbb{I}(.)$ is the indicator function.

Computational complexity of solving the DP is polynomial in τ, but exponential in M. This complexity grows even faster as $N > 2$. Hence, calculating the optimal solution is challenging. However, we will propose a fast greedy heuristic that approximates the optimal solution well.

A Greedy Heuristic

The greedy heuristic (Algorithm 2) essentially ignores time, and at each time-slot sends a subset of packets by the AP's which would maximize the expected number of deliveries for that specific time-slot. Moreover, according to Lemma 23, for finding the subset of packets which results in the maximum expected delivery for a time-slot, it is not necessary to search over all N^M possible subsets; instead, it is sufficient to only focus on N^N subset of them. Algorithm 2 is repeated for all intervals.

In fact, $\sum_{i=1}^{M}(1 - \Pi_{j_m=i,1\leq m\leq N}(1 - p_{mi}))\}$ is the expected number of deliveries for a time-slot, when j_m is transmitted by AP_m. The following lemma establishes why if packets are ordered in the queues of AP's, then for finding the subset of packets which results in maximum expected delivery for the time-slot, it is sufficient to just look at the first N elements of each queue.

Algorithm 2:

Set $t = 1$ and $U = \{1, 2, \ldots, M\}$.

Create N vectors $L_1, \ldots, L_N$, and put the packets $\{1, 2, \ldots, M\}$ in all of them.

Order packets in each L_k, $k \in \{1, 2, \ldots, N\}$, according to p_{kj}'s and in decreasing order.

while $t \leq \tau, U \neq \Phi$ **do**

Find $[j_1, \ldots, j_N] = \arg\max_{j_1 \in L_1(1:N), \ldots, j_N \in L_N(1:N)}$
$\{\sum_{i=1}^{M}(1 - \Pi_{j_m = i, 1 \leq m \leq N}(1 - p_{mi}))\}$.

Transmit $j_1, j_2, \ldots, j_N$ by $\mathsf{AP}_1, \mathsf{AP}_2, \ldots, \mathsf{AP}_N$ respectively.

Update $L_1, L_2, \ldots, L_N$ according to the outcome of transmissions (remove any of $j_1, j_2, \ldots, j_N$ from them which is successfully delivered, and shift the queues to the left to fill the gap of the removed packets). Also, remove the delivered packets from U.

$t \leftarrow t + 1$

end while

Lemma 23. *Suppose*

$$[j_1, j_2, \ldots, j_N] = \arg \max_{j_1 \in L_1(1:N), \ldots, j_N \in L_1(1:N)} \{\sum_{i=1}^{M}(1 - \Pi_{j_m = i, 1 \leq m \leq N}(1 - p_{mi}))\}.$$

If $1 \geq p_{kL_k(1)} \geq p_{kL_k(2)} \geq \ldots, \geq p_{kL_k(|U|)} \geq 0$ *for* $k \in \{1, 2, \ldots, N\}$, *then,*

$$\sum_{i=1}^{M}(1 - \Pi_{j_m = i, 1 \leq m \leq N}(1 - p_{mi})) = \max_{j'_1, \ldots, j'_N \in \{1, 2, \ldots, M\}} \{\sum_{i=1}^{M}(1 - \Pi_{j'_m = i, 1 \leq m \leq N}(1 - p_{mi}))\}.$$

Proof. Consider the N vectors $L_1, \ldots, L_N$, defined in Algorithm 2, where packets in each L_k, $k \in \{1, 2, \ldots, N\}$ are ordered according to p_{kj}'s and in decreasing order, meaning that $1 \geq p_{kL_k(1)} \geq p_{kL_k(2)} \geq \ldots, \geq p_{kL_k(M)} \geq 0$ for $k \in \{1, 2, \ldots, N\}$. Suppose there is no subset of packets $j_1, j_2, \ldots, j_N$ such that each j_k is one of the first N elements of L_k, and $j_1, j_2, \ldots, j_N$ maximizes the expected deliveries over a time-slot for the set of packets $\{1, 2, \ldots, M\}$. More precisely, suppose there is no $j_1, j_2, \ldots, j_N$ such that $j_k \in L_k(1 : N)$ for $k \in \{1, 2, \ldots, N\}$, and it maximizes the $\sum_{i=1}^{M}(1 - \Pi_{j_m=i, 1\leq m\leq N}(1 - p_{mi}))\}$. Consider an arbitrary $j_1, j_2, \ldots, j_N$ which maximizes the expected delivery $\sum_{i=1}^{M}(1 - \Pi_{j_m=i, 1\leq m\leq N}(1 - p_{mi}))\}$. Therefore, there is one of the j_k's that does not belong to the first N elements of L_k. More precisely, there exists a k, $k \in \{1, 2, \ldots, N\}$, for which $j_k \notin L_k(1 : N)$. Therefore, there is at least one of the first N elements of L_k which is not going to be transmitted by any AP. In other words, there must exist an l such that $l \in L_k(1 : N)$, and $l \neq j_i$ for $i \in \{1, 2, \ldots, N\}$. Since, $p_{kl} > p_{kj_k}$, by serving l on AP_k the expected deliveries, $\sum_{i=1}^{M}(1 - \Pi_{j_m=i, 1\leq m\leq N}(1 - p_{mi}))\}$, will only increase. This contradicts the assumption that $j_1, j_2, \ldots, j_N$ produce the maximal expected number of deliveries; and therefore, $j_k \in L_k(1 : N)$ for $k \in \{1, 2, \ldots, N\}$, and the proof is complete. Note that although the lemma and its proof are stated for the set of packets $\{1, 2, \ldots, M\}$, they hold for any arbitrary set of packets U, too. □

The total processing time of Algorithm 2 is $O(\tau MN^{N+1})$; since the while loop is run τ times, and finding $j_1, j_2, \ldots, j_N$ takes $O(N^N MN) = O(MN^{N+1})$.

Numerical Results

We compare the total timely throughput capacity for optimal online policies, splitting policies (C_{T^3}), and greedy heuristic (Algorithm 2).

Heuristic Algorithm 2 is not optimal in general. However, as numerical results indicate, the value provided by the greedy algorithm is quite close to the optimal value. In fact, the numerical results suggest that Algorithm 2 is a decent approximation for the optimal value.

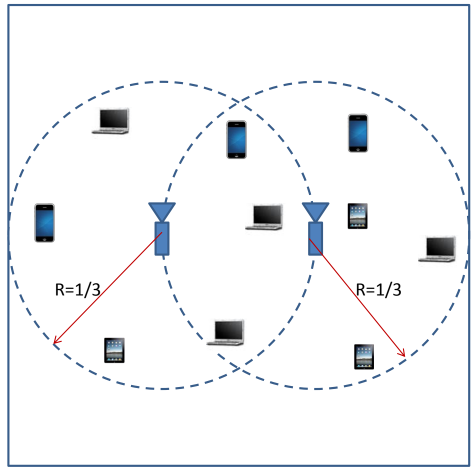
Interestingly, as numerical results in Fig. 5.3(b) indicate, the throughput of offline splitting algorithm is very close to that of optimal online scheduling which is the maximum throughput over all possible policies. Hence, lifting the assumption of partitioning traffic among AP's provides marginal gain over the optimal splitting algorithm, while it requires much more coordination of ACK/NACKs. Consequently, for a system-level design, one may only focus on how to split the traffic among different AP's, and they are ensured that the solution will be near optimal.

5.6.4 Fading Channels and Rate Adaptation

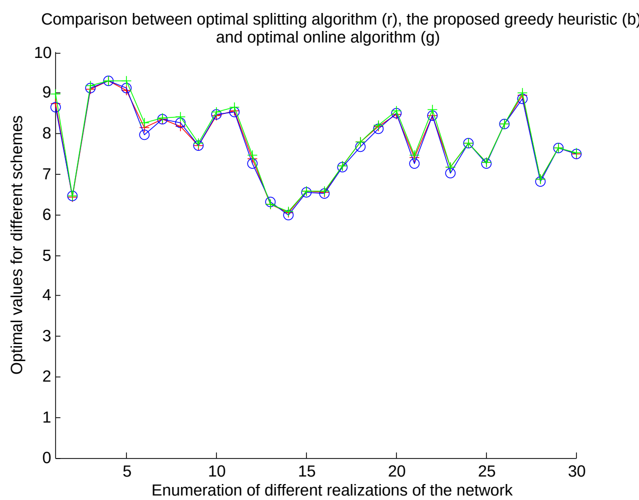
Section 5.2 considered a packet erasure model for channels, and assumed that each AP can transmit one packet at a time. We extend the model to consider fading channels in order to better capture the channel physical properties. In addition, we allow each AP to allocate a portion of its available bandwidth to each client during a time-slot. This means that each AP can access several clients simultaneously. Moreover, we allow for rate adaptation, where according to the time-frequency resource allocated to each client, a certain *reward* will be obtained.

Model Setup

Consider the network topology and time model described in Section 5.2. In addition, for $i \in \{1, 2, \ldots, N\}$, AP_i has bandwidth W_i, where $W_i \in \mathbb{N}$, which means



(a)



(b)

Figure 5.3: (a) wireless network with two AP's with coverage radius $\frac{1}{3}$, 10 randomly and uniformly located clients in the coverage area of the AP's with channel erasure probabilities proportional to the distance, and $\tau = 15$, for 30 different realizations of the network. In (b) the red curve demonstrates the total timely throughput capacity when the scheduling is restricted to partitioning the set of packets across AP's. The green curve demonstrates the total timely throughput capacity when the splitting assumption is relaxed. The blue curve demonstrates the total timely throughput achieved by the greedy heuristic described in Algorithm 2. The curves demonstrate that (i) the greedy heuristic solution is near optimal; and (ii) very marginal capacity gain can be obtained by relaxing the partitioning assumption.

at most W_i simultaneous transmissions can occur during a time-slot by AP_i. On the other hand, all the bandwidth of AP_i during a time-slot might be allocated to a certain client.

Define $R_j^{i_1,i_2,\ldots,i_N}$ to be the total reward obtained by Rx_j during an interval if it is served $i_1, i_2, \ldots, i_N$ times on $\mathsf{AP}_1, \mathsf{AP}_2, \ldots, \mathsf{AP}_N$, respectively. The amount of this reward is determined by the rate adaptation which is used in the AP's. Further, assume that $R_j^{i_1,i_2,\ldots,i_N}$ for $j = 1, 2, \ldots, M$ is a non-negative, increasing function in all dimensions $i_1, i_2, \ldots, i_N$.

A scheduling policy η for the system allocates, possibly at random, the bandwidth of each AP to different clients in each time-slot, based on the past history of the system. Let $q_j(k)$ denote the reward obtained for client Rx_j during interval k under some scheduling policy. The average reward for Rx_j is defined as $q_j = \limsup_{k\to\infty} \frac{\sum_{i=1}^k q_j(i)}{k}$. The objective is to maximize $\sum_{j=1}^M q_j$, which is the total average reward.

Remark 28. *The Relaxed Problem introduced in Section 5.3 was in fact a deterministic scheduling problem with binary rewards; i.e. either size $1/p_{ij}$ would be allocated to packet of client Rx_j in bin i, which would result in reward one (it will contribute to the objective function by setting $x_{ij} = 1$); or, it would not add to the value of the objective function at all (for $x_{ij} = 0$). Therefore, the value of $\sum_{i=1}^N \sum_{j=1}^M x_{ij}$ can be viewed as the reward resulting from a scheduling policy. Nevertheless, a more practical model for the reward is a function with input argument being the amount of time-frequency allocated to the client. Therefore, the model extension we are considering can also be viewed as a generalization of the deterministic scheduling (RP).*

A similar model has been considered in [31] for $N = 1$, where no simultaneous transmissions are allowed, i.e. the bandwidth of AP is equal to 1, and intervals for clients are not necessarily equal. They show that for checking if a set of reward requirements is feasible, it is sufficient to look at the average behavior of the system.However, when going from one AP to multiple AP's checking the average behavior is not sufficient, even when multiple simultaneous transmissions is not allowed, and all deadlines are equal. We focus on maximizing the total average reward, which is the equivalent of C_{T^3} in our original result. To this aim, we first state the following lemma which reduces the problem to a maximization problem over an interval of length τ. Then, we show that this new maximization problem can be solved using Dynamic Programming.

Lemma 24. *The optimization problem can be formulated as follows.*

$$\begin{aligned} \max \quad & \sum_{j=1}^{M} R_j^{x_{1j},\ldots,x_{Nj}} \\ s.t. \quad & \sum_{j=1}^{M} x_{ij} \le W_i \tau, \\ & x_{ij} \in \mathbb{Z}^+ \cup \{0\}, \quad i \in [1:N], j \in [1:M]. \end{aligned} \tag{5.24}$$

Proof. It is sufficient to show that the maximal total average reward is obtained using a policy which is implemented for all intervals; since (5.24) finds the maximal total reward over one interval. The proof in essence is similar to that of Lemma 20. Consider the following two observations. First, we have a finite number of possible actions to take for each interval. More specifically, since we have M clients, N AP's, and $W_i\tau$ chunks of resource in AP_i, total number of different possible actions for an interval is $M^{\sum_{i=1}^{N} \tau W_i}$. Second, each policy produces a certain reward. Among all possible policies for one interval, there is one policy $\mathcal{P}$ with maximal total reward R^*. Hence, any sequence of policies that is implemented on the sequence of intervals produces at most the total average reward of R^*, which is obtained by applying $\mathcal{P}$ to all intervals. □

Dynamic Programming Solution

In this part we use Lemma 24 to propose a DP solution to the problem. Define $OPT[m, t_1, \ldots, t_N]$ to be the maximal total reward obtained when only scheduling the first m clients, with the available resource being $t_1, t_2, \ldots, t_N$ on $\mathsf{AP}_1, \mathsf{AP}_2, \ldots, \mathsf{AP}_N$, respectively. Hence, our objective is to find $OPT[M, W_1\tau, W_2\tau, \ldots, W_N\tau]$.

Algorithm 3:

Input $R_j^{i_1,i_2,\ldots,i_N}$ for $1 \le j \le M, 0 \le i_1 \le \tau W_1$,
$0 \le i_2 \le \tau W_2, \ldots, 0 \le i_N \le \tau W_N$.
Initialize a $[M \times (W_1\tau + 1) \times \ldots \times (W_N\tau + 1)]$ array OPT.
for $m = 1, 2, \ldots, M$ **do**
  for $t_1 \in [0 : \tau W_1], \ldots, t_N \in [0 : \tau W_N]$ **do**
    if $m = 1$ **then**
      $OPT[m, t_1, \ldots, t_N] \leftarrow R_1^{t_1,\ldots,t_N}$;
    else
      $OPT[m, t_1, \ldots, t_N] \leftarrow \max_{0 \le x_1 \le t_1, \ldots, 0 \le x_N \le t_N}$
      $\{OPT[m-1, t_1 - x_1, \ldots, t_N - x_N] + R_m^{x_1,\ldots,x_N}\}$;
    end if
  end for
end for
Output $OPT[M, W_1\tau, \ldots, W_N\tau]$.

Theorem 12. *Algorithm 3 solves the problem of finding the maximum total average reward in $O(M\tau^{2N}\Pi_{i=1}^N W_i^2)$ time.* [1]

[1]The same methodology of applying Dynamic Programming can be used to solve the prob-

Proof. The proof contains two parts: processing time of the algorithm, and proof of correctness.

Total Processing Time: there are total of $O(M\tau^N \Pi_{i=1}^N W_i)$ iterations, each with computational complexity of $O(\tau^N \Pi_{i=1}^N W_i)$. Therefore, the total processing time is $O(M\tau^{2N} \Pi_{i=1}^N W_i^2)$, which is polynomial in the number of clients. (Also, note that the number of AP's, N, is typically small, around 2,3, or 4.)

Proof of Correctness: The algorithm stores an $(N + 1)$-dimensional array OPT. We use induction over the entries of the dynamic programming table, in order that the algorithm fills them in. Induction hypothesis is that $OPT[m, t_1, \ldots, t_N]$ is the maximal total reward when there are only the first m clients in the system, and the available resource on $\mathsf{AP}_1, \ldots, \mathsf{AP}_N$ are $t_1, \ldots, t_N$ respectively. For the base case of $m = 1$ the algorithm allocates all the available resource to the first client, and the table is initialized correctly. We now check the induction step. Consider the time when $OPT[m, t_1, \ldots, t_N]$ is going to be computed by the algorithm; and assume all the previous entries of the table OPT have been correctly computed. First, note that all the entries of the table that the recursive formula for finding $OPT[m, t_1, \ldots, t_N]$ is referring to have already been computed in earlier steps. Second, note that the maximization in the recursive formula accounts for all the possible allocations of the resource to the m-th client, and then for each allocation it computes the maximal total reward, which is the reward using that allocation for client m plus the maximal reward for the subproblem of only having the first $m - 1$ clients, which is already computed correctly according to the induction hypothesis. □

---

lem when packets arrive at the beginning of intervals, but they have different deadlines during the interval.

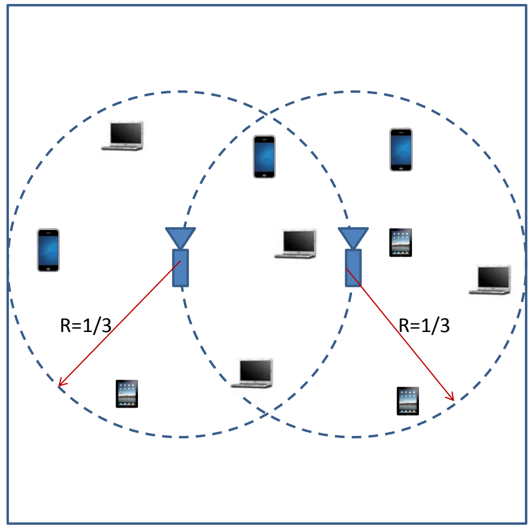
5.7 Numerical Experiments

In this section we provide numerical experiments for our deterministic relaxation scheme. So, we consider a wireless network with 2 AP's, and several wireless clients that are uniformly and randomly located in the network (see Figure 5.4(a)). Channel from every AP to every client is an erasure channel with erasure probability which is proportional to the distance between the AP and the client. The distances in the network are normalized, and we assume that the AP's have the same coverage radius $R = \frac{1}{3}$. Therefore, the channel erasure probability is 1 for the channel between an AP and a client which is located at the distance $R \geq \frac{1}{3}$ from it. Furthermore, the distance between the two AP's is $\frac{1}{3}$.

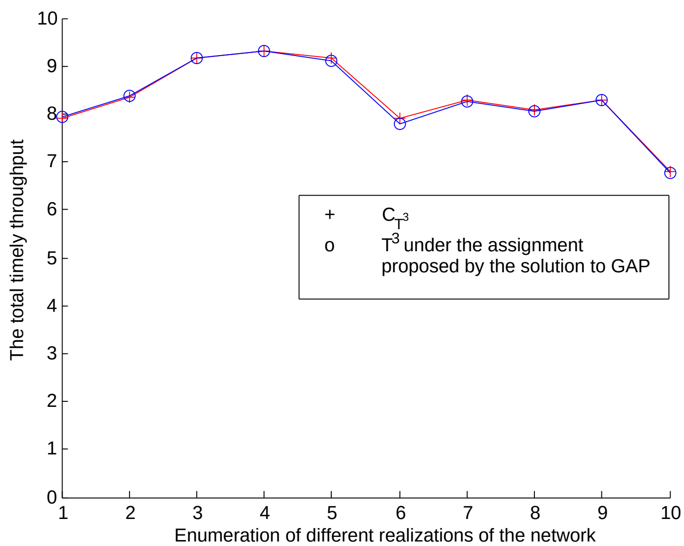
Figure 5.4(b) corresponds to the case where $M = 10$ and $\tau = 15$. In each realization 10 clients are randomly located in the network. For each realization C_{T^3} is calculated. Then, the corresponding relaxed problem is solved, and the network is run for 10000 intervals under the assignments proposed by its deterministic relaxation solution. Fig 5.4(b) shows the comparison between the two for 30 different realizations of the network.

Figure 5.5(a) demonstrates how our proposed LP-rounding algorithm (Algorithm 1) performs compared to C_{det}. We consider $M = 20$ and $\tau = 30$, and 30 different realizations of network. For each realization C_{det}, and the value proposed by our approximation algorithm (Algorithm 1) are found. The result confirms the fact that our proposed algorithm performs well in approximating the optimal solution. The performance improves as the number of clients increases.

Figure 5.5(b) shows how far our T^3 will be from C_{T^3} if we use Algorithm 1 as the assignment strategy for the packets, and run the network for 10000 intervals

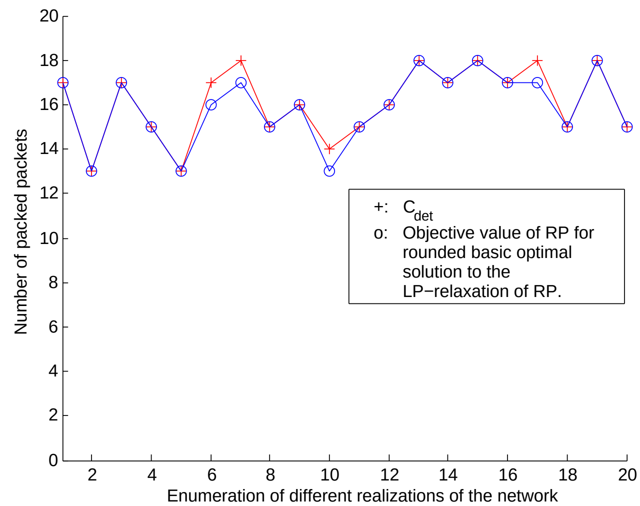


(a)

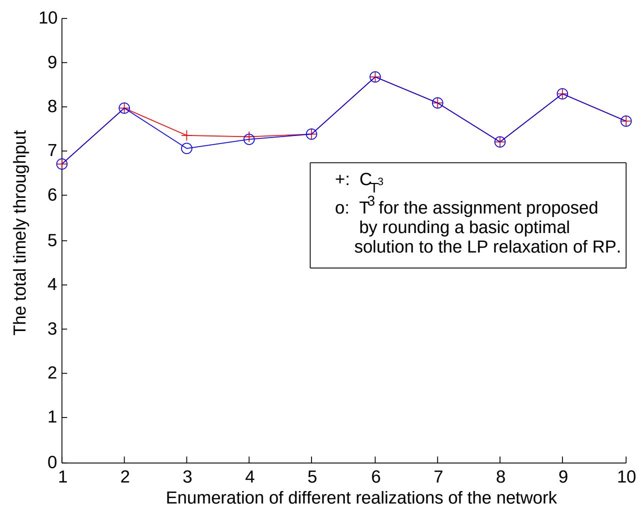


(b)

Figure 5.4: Numerical Results. (a) illustrates the network configuration with two AP's with coverage radius $\frac{1}{3}$ each, M randomly and uniformly located clients in the coverage area of the AP's, and channel erasure probabilities proportional to the distances. (b) compares C_{T^3} with the T^3 resulted from the assignment proposed by (5.4), η_{det}, for $M = 10, \tau = 15$ and 10 different realizations of the network. '+' and 'o' in (b) indicate the values of C_{T^3} and $\mathsf{T}^3(\eta_{\text{det}})$ for each realization, respectively.



(a)



(b)

Figure 5.5: Numerical Results. (a) compares C_{det} (denoted by '+') with the objective value of the rounded basic optimal solution (denoted by 'o') for $M = 20, \tau = 30$ and 20 different network realizations. Finally, (b) compares C_{T^3} (denoted by '+') with the T^3 resulted from the assignment proposed by Algorithm 1 (denoted by 'o') for $M = 10, \tau = 15$ and 10 different realizations of the network.

according to that assignment. In this case we have considered $M = 10, \tau = 15$, and 10 different instances of the network.

5.8 Concluding Remarks and Future Directions

In this chapter we focused on time-sensitive traffic and investigated the improvement by utilizing network heterogeneity in order to enhance the timely throughput of a wireless network. In particular, we studied the problem of maximizing total timely throughput of the downlink of a wireless network with N Access points and M clients, where each client might have access to several Access points. This problem is challenging to attack directly. However, we proposed a deterministic relaxation of the problem which is based on converting the problem to a network with deterministic delay for each link.

First, we showed that the value of the solution to the relaxed problem, C_{det}, is very close to the value of the solution to the original problem, C_{T^3}. In fact, as $C_{\mathrm{T}^3} \to \infty, \frac{C_{\text{det}}}{C_{\mathrm{T}^3}} \to 1$. Furthermore, the numerical results indicate that for networks with limited number of clients, the gap between C_{T^3} and C_{det} is very small. Second, we proposed a simple polynomial-time algorithm with additive performance guarantee of N for approximating the relaxed problem. This approximation performs well as N is for most cases between 2-4. We also extended the formulation to allow time-varying channels, real-time traffic, and weighted total timely throughput maximization, and proved similar results. In addition, we extended the model to account for fading, multiple simultaneous transmissions by Access Points, and rate adaptation. Two future directions are considering multi-hop model, and allowing different deadlines for clients.

CHAPTER 6

CONCLUSION

With the exponential growth in the traffic volume over wireless networks, there is a pressing need for new mechanisms that improve Quality of Service (QoS) in wireless networks. In order to improve QoS, one first needs to understand two important aspects of the problem: (i) requirements and characteristics of traffic flow, and (ii) characteristics of communication network.

A growing portion of the traffic is being occupied by time-sensitive applications which require strict-per-packet deadlines. On the other hand, there are important features of wireless networks, such as channel state information available at the transmitters (CSIT), that can help improve communication rate. However, CSIT can be noisy, delayed, or non-existent, and different receivers might supply CSIT of different qualities. Moreover, networks are evolving towards heterogeneous structures which provide opportunities, as well as new challenges. In this dissertation we developed novel tools and new communication mechanisms that allow for a better understanding of impacts of delayed CSIT, heterogeneous CSIT, and deadline on fundamental limits of communications in modern wireless networks.

In Chapter 2, we focused on the impact of delayed CSIT on fundamental limits of communication over wireless networks. Although there have been two important converse techniques (genie-aided channel enhancement [65] and statistical equivalence of channel outputs [87]) developed in the literature to capture the impact of delayed CSIT, there is still a clear lack of understanding, and the existing techniques fail to address a broad range of networks under delayed CSIT. We developed a novel tool named Rank Ratio Inequality, which captures the impact of delayed CSIT and distributed antennas. This lemma and its proof provide new insights on fundamental limits of signaling in networks with delayed CSIT, and have broad applications for a variety of networks, as presented in Chapter 2 and Chapter 3. Moreover, in Chapter 2 we studied interference channel with delayed CSIT, and characterized the impact of transmitter cooperation on degrees of freedom (DoF). We also studied the MISO broadcast channel with delayed CSIT and presented a new achievable scheme which strictly improves the state-of-the-art scheme.

In Chapter 3 we focused on information-theoretic security, and in particular the problem of wiretap channel, where channels are time-varying and only delayed CSIT is available with respect to the legitimate receiver. This problem has only been solved for the case where eavesdroppers supply delayed CSIT as well. We considered the setting where there is no eavesdropper CSIT, and used several key techniques to completely characterize the secure degrees of freedom (SDoF) of wiretap channel when transmission aided by a distributed jammer which helps jam the confidential message at the eavesdroppers, and when all nodes in the network are equipped with multiple antennas.

In Chapter 4 we studied the impact of heterogeneous CSIT in the context of MISO broadcast channel (MISO BC), where each receiver supplies perfect, delayed, or no CSIT. This problem is quite challenging to the extent that only the DoF for 2-user MISO BC is characterized [20, 81]. We provided a new mathematical tool, called Interference Decomposition Bound, which allows for complete characterization of linear DoF of 3-user MISO BC with heterogeneous CSIT, and also allows for a scalable analysis that can be extended to the general k-user setting.

Finally, in Chapter 5 we considered the evolution of wireless networks into heterogeneous networks, where each user is equipped with multiple wireless access technologies and can receive data from a variety of Access Points. In particular, we considered the downlink of a wireless network with N Access Points (AP's) and M clients, where each client is connected to several out-of-band AP's, and requests delay-sensitive traffic (e.g., real-time video). We characterized the total timely throughput of heterogeneous wireless networks, and extended the result in several directions. The developed tools and insights presented in this dissertation allow for better understanding of impacts of delayed CSIT, heterogeneous CSIT, and deadline on fundamental limits of communications in modern wireless networks.

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 Proof of Lemma 5

A.1.1 Proof of $\mathrm{rank}[\mathbf{G}_{11}^n\mathbf{V}_1^n \quad \mathbf{G}_{12}^n\mathbf{V}_2^n] - \mathrm{rank}[\mathbf{G}_{21}^n\mathbf{V}_1^n \quad \mathbf{G}_{22}^n\mathbf{V}_2^n] \overset{a.s.}{\leq} \mathrm{rank}[\mathbf{G}_{11}^{\mathcal{T}}\mathbf{V}_1^{\mathcal{T}} \quad \mathbf{G}_{12}^{\mathcal{T}}\mathbf{V}_2^{\mathcal{T}}]$:

For a fixed linear coding strategy $\{f_1^{(n)}, f_2^{(n)}\}$, with corresponding $\mathbf{V}_1^n, \mathbf{V}_2^n$, let $\mathcal{A}_i, \mathcal{B}_i, C_i$, $i = 1, 2, \ldots, n$, denote the following sets:

- $\mathcal{A}_i \triangleq \{\mathcal{G}^n| \quad \mathrm{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] = \mathrm{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]\}$.
- $\mathcal{B}_i \triangleq \{\mathcal{G}^n| \quad [\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}], [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \in \mathrm{rowspan}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]\}$.
- $C_i \triangleq \{\mathcal{G}^n| \quad \mathrm{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] = \mathrm{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1\}$.

Note that $\mathcal{B}_i$ is equivalent to $\{\mathcal{G}^n| \quad i \in \mathcal{T}(\mathcal{G}^n)\}$. In order to prove Lemma 5 we first state the following lemma, whose proof is postponed to Appendix A.2.

Lemma 25.

$$\Pr(\boldsymbol{\mathcal{G}}^n \in \cup_{i=1}^n (\mathcal{A}_i \cap \mathcal{B}_i^c)) = 0. \tag{A.1}$$

Lemma 25 implies that we need to prove the first inequality in Lemma 5 only for channel realizations $\boldsymbol{\mathcal{G}}^n = \mathcal{G}^n$, such that $\mathcal{G}^n \notin \cup_{i=1}^n (\mathcal{A}_i \cap \mathcal{B}_i^c)$ (since, the rest have probability measure zero). Thus, we only need to show that for any arbitrary channel realization $\boldsymbol{\mathcal{G}}^n = \mathcal{G}^n$ with the corresponding beamforming matrices V_1^n, V_2^n, and $\boldsymbol{\mathcal{T}} = \mathcal{T}$, such that $\mathcal{G}^n \notin \cup_{i=1}^n (\mathcal{A}_i \cap \mathcal{B}_i^c)$, we have

$$\mathrm{rank}[G_{11}^n V_1^n \quad G_{12}^n V_2^n] - \mathrm{rank}[G_{21}^n V_1^n \quad G_{22}^n V_2^n] \leq \mathrm{rank}[G_{11}^{\mathcal{T}} V_1^{\mathcal{T}} \quad G_{12}^{\mathcal{T}} V_2^{\mathcal{T}}]. \tag{A.2}$$

Let $I(.)$ denote the indicator function, we now bound the left hand side of (A.2) as follows.

$$
\begin{aligned}
&\text{rank}[G_{11}^n V_1^n \quad G_{12}^n V_2^n] - \text{rank}[G_{21}^n V_1^n \quad G_{22}^n V_2^n] \\
&= \sum_{i=1}^{n} (\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] - \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}]) \\
&\quad - (\text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] - \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]) \\
&\leq \sum_{i=1}^{n} \max\{0, (\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] - \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}]) \\
&\quad - (\text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] - \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}])\} \\
&\overset{(a)}{=} \sum_{i=1}^{n} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] = \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1) \\
&\quad \times I(\text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] = \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]) \\
&= \sum_{i=1}^{n} I(\mathcal{G}^n \in \mathcal{A}_i \cap C_i) \\
&= \sum_{i=1}^{n} (I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i \cap C_i) + I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c \cap C_i)) \\
&\leq \sum_{i=1}^{n} (I(\mathcal{G}^n \in \mathcal{B}_i \cap C_i) + I(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c)) \\
&\overset{(b)}{=} \sum_{i=1}^{n} I(\mathcal{G}^n \in \mathcal{B}_i \cap C_i) \overset{(c)}{=} \sum_{i \in \mathcal{T}} I(\mathcal{G}^n \in C_i) \\
&= \sum_{i \in \mathcal{T}} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] = \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1), \qquad \text{(A.3)}
\end{aligned}
$$

where (a) holds since $\text{rank}[G_{k1}^i V_1^i \quad G_{k2}^i V_2^i] - \text{rank}[G_{k1}^{i-1} V_1^{i-1} \quad G_{k2}^{i-1} V_2^{i-1}] \in \{0, 1\}$ for $k = 1, 2$; and (b) follows from the assumption that $\mathcal{G}^n \notin (\mathcal{A}_i \cap \mathcal{B}_i^c)$ for $i \in \{1, 2, \ldots, n\}$; and (c) follows from the fact that $\mathcal{T} = \{i | \mathcal{G}^n \in \mathcal{B}_i\}$. We now only need to show the following to complete the proof of (A.2).

$$
\sum_{i \in \mathcal{T}} I(\text{rank}[G_{11}^i V_1^i \quad G_{12}^i V_2^i] = \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1) \leq \text{rank}[G_{11}^{\mathcal{T}} V_1^{\mathcal{T}} \quad G_{12}^{\mathcal{T}} V_2^{\mathcal{T}}]. \qquad \text{(A.4)}
$$

Without loss of generality, let us assume that $\mathcal{T} = \{\tau_1, \tau_2, \ldots, \tau_k\}$ for some k, such that $\tau_1 < \tau_2 < \ldots < \tau_k$. We define $\mathcal{T}_j \triangleq \{\tau_1, \tau_2, \ldots, \tau_j\}$, and use $V_1^{\mathcal{T}_j}$ and $V_2^{\mathcal{T}_j}$ to denote the sub-matrices of V_1^n and V_2^n with rows in $\mathcal{T}_j$. We also use $G_{11}^{\mathcal{T}_j}$ to denote the $|\mathcal{T}_j| \times |\mathcal{T}_j|$ diagonal matrix with channel coefficients of $g_{11}(t)$ at timeslots $t \in \mathcal{T}_j$ on its diagonal (similarly defined for other channel matrices). We now present a claim that will be used to show (A.4) and complete the proof.

Claim 4. *For any* $j = 1, 2, \ldots, k$,

$$\begin{aligned} &I(\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] = \text{rank}[G_{11}^{\tau_j - 1} V_1^{\tau_j - 1} \quad G_{12}^{\tau_j - 1} V_2^{\tau_j - 1}] + 1) \\ &\leq I(\text{rank}[G_{11}^{\mathcal{T}_j} V_1^{\mathcal{T}_j} \quad G_{12}^{\mathcal{T}_j} V_2^{\mathcal{T}_j}] = \text{rank}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}] + 1). \end{aligned} \tag{A.5}$$

Proof. The claim is trivially true when

$$\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] = \text{rank}[G_{11}^{\tau_j - 1} V_1^{\tau_j - 1} \quad G_{12}^{\tau_j - 1} V_2^{\tau_j - 1}].$$

So, suppose $\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] = \text{rank}[G_{11}^{\tau_j - 1} V_1^{\tau_j - 1} \quad G_{12}^{\tau_j - 1} V_2^{\tau_j - 1}] + 1$. It means that $[g_{11}(\tau_j) \vec{v}_1(\tau_j)^\top \quad g_{12}(\tau_j) \vec{v}_2(\tau_j)^\top]$ is linearly independent of $\text{rowspan}[G_{11}^{\tau_j - 1} V_1^{\tau_j - 1} \quad G_{12}^{\tau_j - 1} V_2^{\tau_j - 1}]$. Since $\mathcal{T}_{j-1} \subseteq \{1, 2, \ldots, \tau_j - 1\}$, then $[g_{11}(\tau_j) \vec{v}_1(\tau_j)^\top \quad g_{12}(\tau_j) \vec{v}_2(\tau_j)^\top]$ is also linearly independent of $\text{rowspan}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}]$. Hence,

$$\text{rank}[G_{11}^{\mathcal{T}_j} V_1^{\mathcal{T}_j} \quad G_{12}^{\mathcal{T}_j} V_2^{\mathcal{T}_j}] = \text{rank}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}] + 1.$$

□

Based on this claim, the proof of (A.4) is as follows.

$$\begin{aligned} &\sum_{i \in \mathcal{T}} I(\text{rank}[G_{11}^{i} V_1^{i} \quad G_{12}^{i} V_2^{i}] = \text{rank}[G_{11}^{i-1} V_1^{i-1} \quad G_{12}^{i-1} V_2^{i-1}] + 1) \\ &= \sum_{j=1}^{k} I(\text{rank}[G_{11}^{\tau_j} V_1^{\tau_j} \quad G_{12}^{\tau_j} V_2^{\tau_j}] = \text{rank}[G_{11}^{\tau_j - 1} V_1^{\tau_j - 1} \quad G_{12}^{\tau_j - 1} V_2^{\tau_j - 1}] + 1) \end{aligned}$$

$$\overset{\text{Claim 4}}{\leq} \sum_{j=1}^{k} I(\text{rank}[G_{11}^{\mathcal{T}_j} V_1^{\mathcal{T}_j} \quad G_{12}^{\mathcal{T}_j} V_2^{\mathcal{T}_j}] = \text{rank}[G_{11}^{\mathcal{T}_{j-1}} V_1^{\mathcal{T}_{j-1}} \quad G_{12}^{\mathcal{T}_{j-1}} V_2^{\mathcal{T}_{j-1}}] + 1)$$

$$= \text{rank}[G_{11}^{\mathcal{T}_k} V_1^{\mathcal{T}_k} \quad G_{12}^{\mathcal{T}_k} V_2^{\mathcal{T}_k}] = \text{rank}[G_{11}^{\mathcal{T}} V_1^{\mathcal{T}} \quad G_{12}^{\mathcal{T}} V_2^{\mathcal{T}}].$$

A.1.2 Proof of $\text{rank}[\mathbf{V}_j^{\mathcal{T}}] \leq \mathbf{r}_j, \quad (j = 1, 2)$:

It is sufficient to prove that $\text{rank}[\mathbf{V}_1^{\mathcal{T}}] \leq \mathbf{r}_1$, since the other inequality (i.e. $\text{rank}[\mathbf{V}_2^{\mathcal{T}}] \leq \mathbf{r}_2$) can be proven similarly. We show that for any realization $\mathcal{G}^n = \{G_{kj}^n\}_{k,j\in\{1,2\}}$ with the corresponding values $\mathcal{T}$, r_1, and matrices V_1^n, V_2^n, we have $\text{rank}[V_1^{\mathcal{T}}] \leq r_1$. But according to definition of r_1, it is sufficient to prove

$$\text{rowspan}[V_1^{\mathcal{T}}] \subseteq \text{span}\left(\vec{s}_{m_1(n)\times 1} | \quad \exists \vec{l}_{n\times 1} \quad s.t. \quad [\vec{s}^{\top} \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^{\top}[G_{21}^n V_1^n \quad G_{22}^n V_2^n]\right). \tag{A.6}$$

The following proves (A.6), thereby completing the proof for $\text{rank}[V_1^{\mathcal{T}}] \leq r_1$:

$$\begin{aligned}
\text{rowspan}[V_1^{\mathcal{T}}] &= \text{span}(\vec{v}_1(i) | 1 \leq i \leq n, \quad [\vec{v}_1(i)^{\top} \quad \vec{0}_{1\times m_2(n)}], \\
&\qquad [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^{\top}] \in \text{rowspan}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]) \\
&\subseteq \text{span}(\vec{v}_1(i) | 1 \leq i \leq n, \quad [\vec{v}_1(i)^{\top} \quad \vec{0}_{1\times m_2(n)}], \\
&\qquad [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^{\top}] \in \text{rowspan}[G_{21}^{n} V_1^{n} \quad G_{22}^{n} V_2^{n}]) \\
&\subseteq \text{span}\left(\vec{v}_1(i) | 1 \leq i \leq n, \quad [\vec{v}_1(i)^{\top} \quad \vec{0}_{1\times m_2(n)}] \in \text{rowspan}[G_{21}^{n} V_1^{n} \quad G_{22}^{n} V_2^{n}]\right) \\
&\subseteq \text{span}\left(\vec{s}_{m_1(n)\times 1} | \quad \exists \vec{l}_{n\times 1} \quad s.t. \quad [\vec{s}^{\top} \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^{\top}[G_{21}^n V_1^n \quad G_{22}^n V_2^n]\right).
\end{aligned}$$

A.1.3 Proof of $\mathbf{r}_j \overset{a.s.}{\leq} \operatorname{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n] - \operatorname{rank}[\mathbf{V}_{3-j}^n], \quad (j = 1, 2)$:

We will show this for $j = 1$, i.e., $\mathbf{r}_1 \overset{a.s.}{\leq} \operatorname{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n] - \operatorname{rank}[\mathbf{V}_2^n]$. The proof for $j = 2$ will be similar. Since $\operatorname{rank}[\mathbf{G}_{22}^n \mathbf{V}_2^n] \overset{a.s.}{=} \operatorname{rank}[\mathbf{V}_2^n]$, it is sufficient to show that $\mathbf{r}_1 \leq \operatorname{rank}[\mathbf{G}_{21}^n \mathbf{V}_1^n \quad \mathbf{G}_{22}^n \mathbf{V}_2^n] - \operatorname{rank}[\mathbf{G}_{22}^n \mathbf{V}_2^n]$. To do so, we show that for any realization $\mathcal{G}^n = \{G_{kj}^n\}_{k,j\in\{1,2\}}$ with the corresponding value r_1, and matrices V_1^n, V_2^n, we have $r_1 \leq \operatorname{rank}[G_{21}^n V_1^n \quad G_{22}^n V_2^n] - \operatorname{rank}[G_{22}^n V_2^n]$.

Since $r_1 = \dim(\operatorname{span}(\vec{s}_{m_1(n)\times 1} | \quad \exists \vec{l}_{n\times 1} \quad s.t. \quad [\vec{s}^\top \quad \vec{0}_{1\times m_2(n)}] = \vec{l}^\top [G_{21}^n V_1^n \quad G_{22}^n V_2^n]))$, we have

$$\exists L_{r_1\times n} \quad s.t. \quad [S \quad 0_{r_1\times m_2(n)}] = L[G_{21}^n V_1^n \quad G_{22}^n V_2^n], \tag{A.7}$$

for some $S_{r_1\times m_1(n)}$, such that $\operatorname{rank}[S] = r_1$. This means

$$LG_{22}^n V_2^n = 0_{r_1\times m_2(n)}, \quad LG_{21}^n V_1^n = S, \quad \operatorname{rank}[S] = r_1. \tag{A.8}$$

We now state a claim that will be useful in completing the proof.

Claim 5. *For three matrices A, B, C where the number of columns in A is equal to the number of rows in B, C,*

$$\operatorname{rank}[AB \quad AC] - \operatorname{rank}[AC] \leq \operatorname{rank}[B \quad C] - \operatorname{rank}[C]. \tag{A.9}$$

Proof. By Frobenius's inequality, for any three matrices X, Y, Z where XY, YZ, and XYZ are defined,

$$\operatorname{rank}[XY] + \operatorname{rank}[YZ] \leq \operatorname{rank}[XYZ] + \operatorname{rank}[Y]. \tag{A.10}$$

By setting $X = A, Y = [B \quad C], Z = [0 \quad I]^\top$, where I is the identity matrix, the desired result follows. $\square$

Therefore, by setting $A = L, B = [G_{21}^n V_1^n \quad G_{22}^n V_2^n], C = G_{22}^n V_2^n$ in Claim 5, and using (A.8), we get

$$r_1 - 0 \leq \text{rank}[G_{21}^n V_1^n \quad G_{22}^n V_2^n] - \text{rank}[G_{22}^n V_2^n], \tag{A.11}$$

which completes the proof.

A.2 Proof of Lemma 25

Here we restate Lemma 25 before proving it.

Lemma 25. *Consider a fixed linear coding strategy* $\{f_1^{(n)}, f_2^{(n)}\}$*, with corresponding* $\mathbf{V}_1^n \triangleq \mathbf{V}_{11}^n, \mathbf{V}_2^n \triangleq \mathbf{V}_{12}^n$ *as defined in (4.4). For any* $i \in \{1, 2, \ldots, n\}$*, let* $\mathcal{A}_i, \mathcal{B}_i$*, denote the following sets:*

- $\mathcal{A}_i \triangleq \{\mathcal{G}^n | \quad \text{rank}[G_{21}^i V_1^i \quad G_{22}^i V_2^i] = \text{rank}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]\}$.
- $\mathcal{B}_i \triangleq \{\mathcal{G}^n | \quad [\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}], [\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \in \text{rowspan}[G_{21}^{i-1} V_1^{i-1} \quad G_{22}^{i-1} V_2^{i-1}]\}$.

Then,

$$\Pr(\mathcal{G}^n \in \cup_{i=1}^n (\mathcal{A}_i \cap \mathcal{B}_i^c)) = 0.$$

Proof. Note that due to Union Bound, it is sufficient to show that for any $i \in \{1, 2, \ldots, n\}$,

$$\Pr(\mathcal{G}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c) = 0.$$

Consider an arbitrary $i \in \{1, 2, \ldots, n\}$. Due to Total Probability Law, it is sufficient to show that for any channel realization of the first $i-1$ timeslots, denoted by

$\mathcal{G}^{i-1} = \{G_{kj}^{i-1}\}_{j,k\in\{1,2\}}$, we have

$$\Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1}) = 0. \tag{A.12}$$

Consider an arbitrary channel realization of the first $i-1$ timeslots $\mathcal{G}^{i-1} = \{G_{kj}^{i-1}\}_{j,k\in\{1,2\}}$, with corresponding matrices V_1^i, V_2^i (which are now deterministic). Also, suppose that given $\mathcal{G}^{i-1}$, $\mathcal{B}_i^c$ occurs; since otherwise, the proof would be complete. On the other hand, assuming $\mathcal{B}_i^c$ occurs, and denoting $\mathcal{L} = \text{rowspan}[G_{21}^{i-1}V_1^{i-1} \quad G_{22}^{i-1}V_2^{i-1}]$, at least one of the following is true according to the definition of $\mathcal{B}_i$:

$$[\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}] \notin \mathcal{L} \qquad \Rightarrow \qquad \text{Proj}_{\mathcal{L}^c}[\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}] \neq 0, \tag{A.13}$$

$$[\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \notin \mathcal{L} \qquad \Rightarrow \qquad \text{Proj}_{\mathcal{L}^c}[\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top] \neq 0. \tag{A.14}$$

Therefore, the $(m_1(n)+m_2(n))\times 2$ matrix

$$[\text{Proj}_{\mathcal{L}^c}[\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}]^\top \qquad \text{Proj}_{\mathcal{L}^c}[\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top]^\top]$$

is non-zero, which means that its null space has dimension strictly lower than 2. Hence, we have,

$$\begin{aligned}
&\Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_i \cap \mathcal{B}_i^c | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1}) \\
&\stackrel{(a)}{=} \Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_i | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1}) \\
&\stackrel{(b)}{=} \Pr(\text{Proj}_{\mathcal{L}^c}[\mathbf{g}_{21}(i)\vec{v}_1(i)^\top \quad \mathbf{g}_{22}(i)\vec{v}_2(i)^\top] = 0 | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1}) \\
&\stackrel{(c)}{=} \Pr(\mathbf{g}_{21}(i)\text{Proj}_{\mathcal{L}^c}[\vec{v}_1(i)^\top \quad \vec{0}] + \mathbf{g}_{22}(i)\text{Proj}_{\mathcal{L}^c}[\vec{0} \quad \vec{v}_2(i)^\top] = 0 | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1}) \\
&= \Pr([\text{Proj}_{\mathcal{L}^c}[\vec{v}_1(i)^\top \quad \vec{0}]^\top \quad \text{Proj}_{\mathcal{L}^c}[\vec{0} \quad \vec{v}_2(i)^\top]^\top] \begin{bmatrix} \mathbf{g}_{21}(i) \\ \mathbf{g}_{22}(i) \end{bmatrix} = 0 | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1}) \\
&= \Pr(\begin{bmatrix} \mathbf{g}_{21}(i) \\ \mathbf{g}_{22}(i) \end{bmatrix} \in \text{nullspace}[\text{Proj}_{\mathcal{L}^c}[\vec{v}_1(i)^\top \quad \vec{0}]^\top \qquad \text{Proj}_{\mathcal{L}^c}[\vec{0} \quad \vec{v}_2(i)^\top]^\top] | \boldsymbol{\mathcal{G}}^{i-1} = \mathcal{G}^{i-1})
\end{aligned}$$

$$\stackrel{(d)}{=} 0,$$

where (a) holds since we assumed that for realization $\mathcal{G}^{i-1}$, $\mathcal{B}_i^c$ occurs; (b) holds according to the definition of $\mathcal{A}_i$; (c) holds due to linearity of orthogonal projection; and (d) holds since the $(m_1(n) + m_2(n)) \times 2$ matrix $[\mathrm{Proj}_{\mathcal{L}^c}[\vec{v}_1(i)^\top \quad \vec{0}_{1\times m_2(n)}]^\top \qquad \mathrm{Proj}_{\mathcal{L}^c}[\vec{0}_{1\times m_1(n)} \quad \vec{v}_2(i)^\top]^\top]$ is non-zero, which means that its null space, which is a subspace in $\mathbb{R}^2$, has dimension strictly lower than 2. Therefore, the probability that the random vector $\begin{bmatrix} \mathbf{g}_{21}(i) \\ \mathbf{g}_{22}(i) \end{bmatrix}$ lies in a subspace in $\mathbb{R}^2$ of strictly lower dimension (than 2) is zero.

$\square$

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 Proof of Least Alignment Lemma (Lemma 9)

We first restate Lemma 9 here.

Lemma 9 (Least Alignment Lemma). *Consider two receivers* $\mathrm{Rx}_1, \mathrm{Rx}_2$ *with* n_0 *antennas, where* Rx_2 *supplies no CSIT. Then, for a given* $n \in \mathbb{N}$ *and any encoding strategy* $f^{(n)}$ *as defined in Definition 6,*

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq h(\vec{y}_2^n|\mathcal{G}^n) + n.o(\log p).$$

We prove a stronger version of Lemma 9. More specifically, we prove that if Tx only knows a probability distribution for values of the channels to Rx_2, and we denote the maximum value of such distribution by $f_{max}(p)$, then

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq h(\vec{y}_2^n|\mathcal{G}^n) + nn_0 \log(f_{max}(p)) + n.o(\log p).$$

In order to prove the above inequality, we use the approach in [21] to generalize that result to the MIMO case. In particular, we first transform the network via a deterministic channel model.

B.1.1 Deterministic Channel Model

To prove the lemma, we first discretize the channel to avoid dealing with the impact of additive Gaussian noise. This leads to a deterministic channel model described as follows. For $j = 1, 2$, let

$$\vec{x}(t) = \begin{bmatrix} \bar{x}_1(t) \\ \vdots \\ \bar{x}_m(t) \end{bmatrix}, \qquad \vec{y}_j(t) = \begin{bmatrix} \bar{y}_{j,\{1\}}(t) \\ \vdots \\ \bar{y}_{j,\{n_0\}}(t) \end{bmatrix}, \qquad G_j(t) = \begin{bmatrix} g_{j,\{1,1\}}(t) & \dots & g_{j,\{1,m\}}(t) \\ \vdots & & \\ g_{j,\{n_0,1\}}(t) & \dots & g_{j,\{n_0,m\}}(t) \end{bmatrix}. \tag{B.1}$$

The channel input at time t is denoted by $\vec{x}(t)$, where $\vec{x}(t) \in \{0, 1, \dots, \lceil \sqrt{p} \rceil\}^m$. The channel outputs are defined as

$$\vec{y}_j(t) = \begin{bmatrix} \bar{y}_{j,\{1\}}(t) \\ \vdots \\ \bar{y}_{j,\{n_0\}}(t) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} \lfloor g_{j,\{1,i\}}(t) \bar{x}_i(t) \rfloor \\ \vdots \\ \sum_{i=1}^{m} \lfloor g_{j,\{n_0,i\}}(t) \bar{x}_i(t) \rfloor \end{bmatrix}. \tag{B.2}$$

Lemma 26.

$$\max \lim_{p \to \infty} \frac{h(\vec{y}_1^n | \mathcal{G}^n) - h(\vec{y}_2^n | \mathcal{G}^n)}{\log p} \leq \max \lim_{p \to \infty} \frac{H(\vec{y}_1^n | \mathcal{G}^n) - H(\vec{y}_2^n | \mathcal{G}^n)}{\log p}, \tag{B.3}$$

where the maximum on the left hand side is taken over all possible encoding strategies as defined in Definition 6; and the maximum on the right hand side is taken over all possible encoding schemes for the deterministic channel.

Proof of Lemma 26 follows from similar arguments used to prove that DoF of a network under deterministic channel model is an upper bound on the actual DoF. The proof can be found in prior works, including [14, 21], and hence is omitted for brevity. Lemma 26 suggests that for proving Lemma 9, it is sufficient to prove that under deterministic channel model,

$$H(\vec{y}_1^n | \mathcal{G}^n) - H(\vec{y}_2^n | \mathcal{G}^n) \leq n.o(\log p). \tag{B.4}$$

As a result, our objective henceforth will be to prove (B.4).

B.1.2 Imposing Functional Dependence

We will show in this section that by imposing functional dependence of $\vec{\bar{x}}^n$ on $(\vec{\bar{y}}_1^n, G_1^n)$, we obtain an upper bound on $H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n)$. Define $\mathcal{L}$ as the mapping from $(\vec{\bar{y}}_1^n, G_1^n)$ to $\vec{\bar{x}}^n$:

$$\vec{\bar{x}}^n = \mathcal{L}(\vec{\bar{y}}_1^n, G_1^n). \tag{B.5}$$

This mapping is in general stochastic, and therefore, $\mathcal{L}$ is a random variable. Hence, by conditioning on $\mathcal{L}$ we obtain

$$H(\vec{\bar{y}}_2^n|\mathcal{G}^n) \geq H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L}) \geq \min_L H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L} = L) = H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L} = L_0), \tag{B.6}$$

where[1] $L_0 \triangleq \arg\min_L H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L} = L)$ is a deterministic map. Note that the choice of map does not impact $(\vec{\bar{y}}_1^n, \mathcal{G}^n)$. Hence, using (B.6) we obtain

$$H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n) \leq H(\vec{\bar{y}}_1^n|\mathcal{G}^n, \mathcal{L} = L_0) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L} = L_0). \tag{B.7}$$

Thus, henceforth, we will upper bound (B.7) in order to complete the proof of Lemma 9; and we will assume that

$$\vec{\bar{x}}^n = L_0(\vec{\bar{y}}_1^n, G_1^n). \tag{B.8}$$

Note that the above equation suggests that $\vec{\bar{y}}_2^n$ is fully specified by $(\vec{\bar{y}}_1^n, \mathcal{G}^n)$; hence, we use the following notation for simplicity:

$$\vec{\bar{y}}_2^n(\vec{\bar{y}}_1^n, \mathcal{G}^n) \triangleq \begin{bmatrix} \sum_{i=1}^m \lfloor g_{2,\{1,i\}}(t) L_0(\vec{\bar{y}}_1^n, G_1^n)_i(t) \rfloor \\ \vdots \\ \sum_{i=1}^m \lfloor g_{2,\{n_0,i\}}(t) L_0(\vec{\bar{y}}_1^n, G_1^n)_i(t) \rfloor \end{bmatrix}, \tag{B.9}$$

where $L_0(\vec{\bar{y}}_1^n, G_1^n) = \vec{\bar{x}}^n$, and $L_0(\vec{\bar{y}}_1^n, G_1^n)_i(t) = \bar{x}_i(t)$, which is the i-th element of the vector $\vec{\bar{x}}(t)$.

[1] In cases where $\arg\min_L H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L} = L)$ is not unique, we choose L_0 to be a deterministic mapping that minimizes $H(\vec{\bar{y}}_2^n|\mathcal{G}^n, \mathcal{L} = L)$.

B.1.3 Upper Bounding $H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n)$ via Aligned Image Sets

Note that our goal is to upper bound $H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n)$. This means that we will try to upper bound the difference between received signal dimensions at $\mathrm{Rx}_1, \mathrm{Rx}_2$. This difference grows when more codewords that are received at Rx_1 as different received codewords align perfectly at Rx_2; because in such case, the dimension of received signal at Rx_2 will decrease, leading to an increase in the difference of received signal dimensions at the two receivers. Hence, we will focus on aligned images sets, which are sets of distinct codewords received at Rx_1 that are aligned at Rx_2. More specifically, for a pair $(\vec{\bar{v}}^n, \mathcal{G}^n)$ of received codeword at Rx_2, $\vec{\bar{v}}^n$, and channel coefficients, $\mathcal{G}^n$, we define the corresponding aligned image set as the set of all received signals at Rx_1 which have the same image at Rx_2 as $\vec{\bar{v}}^n$.

Definition 14. (Aligned Image Set)

$$S_{\vec{\bar{v}}^n}(\mathcal{G}^n) \triangleq \left\{ \vec{\bar{y}}_1^n \quad | \quad \vec{\bar{y}}_2^n(\vec{\bar{y}}_1^n, \mathcal{G}^n) = \vec{\bar{v}}^n, \quad t = 1, \ldots, n \right\}.$$

We now upper bound $H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n)$ via analyzing the cardinality of aligned image sets.

$$\begin{aligned} H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n) &\leq H(\vec{\bar{y}}_1^n, \vec{\bar{y}}_2^n|\mathcal{G}^n, L_0) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n, L_0) = H(\vec{\bar{y}}_1^n|\vec{\bar{y}}_2^n, \mathcal{G}^n, L_0) \\ &\overset{(a)}{=} H(\vec{\bar{y}}_1^n| \quad |S_{\vec{\bar{y}}_2^n}(\mathcal{G}^n)|, \vec{\bar{y}}_2^n, \mathcal{G}^n, L_0) \\ &\overset{(b)}{\leq} E \log |S_{\vec{\bar{y}}_2^n}(\mathcal{G}^n)| \\ &\overset{(c)}{\leq} \log E|S_{\vec{\bar{y}}_2^n}(\mathcal{G}^n)|, \end{aligned} \tag{B.10}$$

where (a) holds since $(\vec{\bar{y}}_2^n, \mathcal{G}^n, L_0)$ completely determines the aligned image set $S_{\vec{\bar{y}}_2^n}(\mathcal{G}^n)$; (b) holds since the entropy of $\vec{\bar{y}}_1^n$ is maximized when it has a uniform distribution over all of its possible values, which are determined by $S_{\vec{y}_2^n}(\mathcal{G}^n)$; and (c) holds due to Jensen's inequality.

Hence, we will focus on upper bounding $E|S_{\vec{y}_2^n}(\mathcal{G}^n)|$. To this aim, for a given $(\vec{y}_1^n, \vec{v}^n)$ we first analyze $\Pr(\vec{y}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n))$, which is the probability that the received image of a certain codeword at Rx_1 has the same image at Rx_2 as $\vec{v}^n$.

B.1.4 Bounding the probability of Image Alignment

In this section we will provide an upper bound on $\Pr(\vec{y}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n))$. To this aim, we analyze $\Pr(\vec{y}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n)|G_1^n)$. Let us fix $\vec{y}_1^n, \vec{v}^n, G_1^n$. Note that given $(\vec{y}_1^n, G_1^n)$, $\vec{x}^n$ is determined. Consider the event where $\vec{y}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n)$. This event is equivalent to $\vec{y}_2^n(\vec{y}_1^n, \mathcal{G}^n) = \vec{v}^n$, which in turn by (B.9) is equivalent to the following:

$$\forall t = 1, \ldots, n, \qquad \begin{bmatrix} \sum_{i=1}^m \lfloor g_{2,\{1,i\}}(t) L_0(\vec{y}_1^n, G_1^n)_i(t) \rfloor \\ \vdots \\ \sum_{i=1}^m \lfloor g_{2,\{n_0,i\}}(t) L_0(\vec{y}_1^n, G_1^n)_i(t) \rfloor \end{bmatrix} = \vec{v}(t) = \begin{bmatrix} \bar{v}_1(t) \\ \vdots \\ \bar{v}_{n_0}(t) \end{bmatrix}, \tag{B.11}$$

or equivalently,

$$\forall t = 1, \ldots, n, \; j = 1, \ldots, n_0, \qquad \sum_{i=1}^m \lfloor g_{2,\{j,i\}}(t) \bar{x}_i(t) \rfloor = \bar{v}_j(t). \tag{B.12}$$

Let $i^*(t) = \arg\max_i \bar{x}_i(t)$. As a result, the above event can be re-written as follows:

$$\forall t = 1, \ldots, n, \; j = 1, \ldots, n_0,$$
$$\bar{v}_j(t) - \sum_{i=1, i \neq i^*(t)}^m \lfloor g_{2,\{j,i\}}(t) \bar{x}_i(t) \rfloor \leq g_{2,\{j,i^*(t)\}}(t) \bar{x}_{i^*(t)}(t) \leq \bar{v}_j(t) - \sum_{i=1, i \neq i^*(t)}^m \lfloor g_{2,\{j,i\}}(t) \bar{x}_i(t) \rfloor + 1.$$

Hence, for every t, if $\bar{x}_{i^*(t)}(t) \neq 0$, then for the event (B.12) to occur it is necessary that $g_{2,\{j,i^*(t)\}}(t)$ fall in an interval of length $\frac{1}{\bar{x}_{i^*(t)}(t)}$ for $t = 1, \ldots, n$, $j = 1, \ldots, n$. Also, note that if $\vec{x}(t) = \vec{0}$ (which means $\bar{x}_{i^*(t)}(t) = 0$), then $\vec{y}_1(t) = \vec{y}_2(t) = \vec{0}$. Hence, for all $t = 1, \ldots, n$, where $\vec{\bar{y}}_1(t) \neq \vec{0}$, the probability of occurrence of (B.12) for $t = 1, \ldots, n$, $j = 1, \ldots, n_0$, is at most $f_{max}(p)(\frac{1}{\bar{x}_{i^*(t)}(t)})$. We now further upper bound this quantity. Note that since $\bar{y}_{1,\{j\}}(t) = \sum_{i=1}^{m} \lfloor g_{1,\{j,i\}}(t)\bar{x}_i(t) \rfloor$,

$$|\bar{y}_{1,\{j\}}(t)| \leq \bar{x}_{i^*(t)}(t) \sum_{i=1}^{m} |g_{1,\{j,i\}}(t)| + m.$$

By re-writing the above inequality, when $|\bar{y}_{1,\{j\}}(t)| > m$,

$$\frac{1}{\bar{x}_{i^*(t)}(t)} \leq \frac{\sum_{i=1}^{m} |g_{1,\{j,i\}}(t)|}{|\bar{y}_{1,\{j\}}(t)| - m}.$$

Hence, we have the following upper bound on the probability of occurrence of $\vec{\bar{y}}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n)$:

$$\begin{aligned}
\Pr(\vec{\bar{y}}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n)|G_1^n) &\leq \prod_{\substack{t:\\ \vec{\bar{y}}_1(t)\neq\vec{0}}} \prod_{\substack{j:\\ |\bar{y}_{1,\{j\}}(t)|>m}} \frac{f_{max}(p)(\sum_{i=1}^{m} |g_{1,\{j,i\}}(t)|)}{|\bar{y}_{1,\{j\}}(t)| - m} \\
&\leq \left(\prod_{t=1}^{n}\prod_{j=1}^{n} \max(1, f_{max}(p)(\sum_{i=1}^{m} |g_{1,\{j,i\}}(t)|))\right) \prod_{t=1}^{n}\prod_{j=1}^{n} \frac{1}{\max(1, |\bar{y}_{1,\{j\}}(t)| - m)} \\
&\leq \max(1, f_{max}(p) m d_{max})^{nn_0} \prod_{t=1}^{n}\prod_{j=1}^{n} \frac{1}{\max(1, |\bar{y}_{1,\{j\}}(t)| - m)}. \qquad \text{(B.13)}
\end{aligned}$$

We use the above bound on $\Pr(\vec{\bar{y}}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n)|G_1^n)$ to further upper bound (B.10).

B.1.5 Bounding the Average Size of Aligned Image Sets

For a given $\vec{v}^n$,

$$\begin{aligned}
E[S_{\vec{v}^n}(\mathcal{G}^n)] = E[E[S_{\vec{v}^n}(\mathcal{G}^n)|G_1^n]] &= E[\sum_{\vec{\bar{y}}_1^n} \Pr(\vec{\bar{y}}_1^n \in S_{\vec{v}^n}(\mathcal{G}^n)|G_1^n)] \\
&\overset{(B.13)}{\leq} \max(1, f_{max}(p) m d_{max})^{nn_0} \sum_{\vec{\bar{y}}_1^n} \prod_{t=1}^{n}\prod_{j=1}^{n} \frac{1}{\max(1, |\bar{y}_{1,\{j\}}(t)| - m)} \\
&\overset{(a)}{=} \max(1, f_{max}(p) m d_{max})^{nn_0} \prod_{t=1}^{n}\prod_{j=1}^{n} \sum_{\bar{y}_{1,\{j\}}(t)} \frac{1}{\max(1, |\bar{y}_{1,\{j\}}(t)| - m)}
\end{aligned}$$

$$\overset{(b)}{\leq} \max(1, f_{max}(p)md_{max})^{nn_0} \prod_{t=1}^{n}\prod_{j=1}^{n}\{ \sum_{\substack{\bar{y}_{1,\{j\}}(t):\\ |\bar{y}_{1,\{j\}}(t)|\leq m}} 1 + \sum_{\substack{\bar{y}_{1,\{j\}}(t):\\ m<|\bar{y}_{1,\{j\}}(t)|\leq q}} \frac{1}{|\bar{y}_{1,\{j\}}(t)| - m}\}$$

$$\leq \max(1, f_{max}(p)md_{max})^{nn_0} \prod_{t=1}^{n}\prod_{j=1}^{n}\{\log q + o(\log q)\}$$

$$\leq \max(1, f_{max}(p)md_{max})^{nn_0}(\log q)^{nn_0} + o(\log q)$$

$$\leq \max(1, f_{max}(p)md_{max})^{nn_0}(\log \sqrt{p})^{nn_0} + o(\log p), \tag{B.14}$$

where (a) follows from interchanging sum and product; and (b) follows from the definition $q \triangleq md_{max}\sqrt{p} + m$, and noting that $|\bar{y}_{1,\{j\}}(t)| \leq q$. Hence, by (B.10) and (B.14), we obtain

$$\begin{aligned} H(\vec{\bar{y}}_1^n|\mathcal{G}^n) - H(\vec{\bar{y}}_2^n|\mathcal{G}^n) &\leq \log\left(\max(1, f_{max}(p)md_{max})^{nn_0}(\log \sqrt{p})^{nn_0} + o(\log p)\right) \\ &= nn_0 \log(f_{max}(p)) + o(\log p). \end{aligned} \tag{B.15}$$

Therefore, the proof of Lemma 9 is complete.

Remark 29. *Note that (B.15) in fact proves a stronger statement than Lemma 9. In particular, for any CSIT quality supplied by* Rx_2 *as a function of power* $f_{max}(p)$*, (B.15) implies that*

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq h(\vec{y}_2^n|\mathcal{G}^n) + nn_0\log(f_{max}(p)) + n.o(\log p).$$

Nevertheless, in order to prove Theorem 8, it was sufficient to consider the special case where $f_{max}(p) = o(\log p)$*, which is the case in the statement of Lemma 9.*

B.2 Proof of Proposition 4

Let us first consider a hypothetical receiver Rx_0 with n_1 antennas for which there is no CSIT available to the transmitter. Hence, by Lemma 9,

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq h(\vec{y}_0^n|\mathcal{G}^n) + n.o(\log p). \tag{B.16}$$

Moreover, since $n_1 \geq \bar{n}$ and there is no CSIT with respect to any of $\text{Rx}_0, \text{Rx}_{max,1}$, using Lemma 10,

$$\frac{h(\vec{y}_0^n|\mathcal{G}^n)}{n_1} \leq \frac{h(\vec{y}_{max,1}^n|\mathcal{G}^n)}{\bar{n}} + n.o(\log p). \tag{B.17}$$

Therefore, by combining the inequalities in (B.16)-(B.17) we get

$$h(\vec{y}_1^n|\mathcal{G}^n) \leq \frac{n_1}{\bar{n}} h(\vec{y}_{max,1}^n|\mathcal{G}^n) + n.o(\log p),$$

which completes the proof of Proposition 4.

B.3 Proof of Lemma 11

We first state an extension of Lemma 10, which is useful in proving Lemma 11, and can be proved using the same proof steps as in proof of Lemma 10.

Lemma 27. *Consider receivers* $\text{Rx}_1, \text{Rx}_2, \text{Rx}_3$ *which supply no CSIT, with* n_1, n_2, n_3 *antennas, where* $n_1 \geq n_2$. *Then, for a given* $n \in \mathbb{N}$ *and any encoding strategy* $f^{(n)}$ *as defined in Definition 6,*

$$\frac{h(\vec{y}_1^n|\vec{y}_3^n, \mathcal{G}^n)}{min(m, n_1)} \leq \frac{h(\vec{y}_2^n|\vec{y}_3^n, \mathcal{G}^n)}{min(m, n_2)} + n \times o(\log p).$$

Consider receivers $\text{Rx}_1, \text{Rx}_2, \text{Rx}_3$ with n_1, n_2, n_3 antennas, where $n_1, n_2, n_3 > 0$, and $m \geq n_1 + n_2 + n_3$. Also, suppose that Rx_2, Rx_3 supply no CSIT. Further, let Rx_0 denote a receiver with n_1 antennas supplying no CSIT.

$$\begin{aligned} n_2 \times h(\vec{y}_1^n|\vec{y}_2^n, \vec{y}_3^n, \mathcal{G}^n) &= n_2 \times h(\vec{y}_1^n, \vec{y}_2^n, \vec{y}_3^n|\mathcal{G}^n) - n_2 \times h(\vec{y}_3^n|\mathcal{G}^n) - n_2 \times h(\vec{y}_2^n|\vec{y}_3^n, \mathcal{G}^n) \\ &\overset{(a)}{\leq} n_2 \times h(\vec{y}_0^n, \vec{y}_2^n, \vec{y}_3^n|\mathcal{G}^n) - n_2 \times h(\vec{y}_3^n|\mathcal{G}^n) - n_2 \times h(\vec{y}_2^n|\vec{y}_3^n, \mathcal{G}^n) + n.o(\log p) \\ &= n_2 \times h(\vec{y}_0^n, \vec{y}_2^n|\vec{y}_3^n, \mathcal{G}^n) - n_2 \times h(\vec{y}_2^n|\vec{y}_3^n, \mathcal{G}^n) + n.o(\log p) \end{aligned}$$

$$\overset{\text{Lemma 27}}{\leq} n_2 \times (\frac{n_1 + n_2}{n_2}) \times h(\vec{y}_2^n | \vec{y}_3^n, \mathcal{G}^n) - n_2 \times h(\vec{y}_2^n | \vec{y}_3^n, \mathcal{G}^n) + n.o(\log p)$$

$$\leq n_1 \times h(\vec{y}_2^n | \vec{y}_3^n, \mathcal{G}^n) + n.o(\log p), \tag{B.18}$$

where (a) holds since the virtual receiver incorporating $(\mathrm{Rx}_0, \mathrm{Rx}_2, \mathrm{Rx}_3)$, which has $n_1 + n_2 + n_3$ antennas, supplies no CSIT; and as a result, we can apply Lemma 9 to lower bound $h(\vec{y}_0^n, \vec{y}_2^n, \vec{y}_3^n | \mathcal{G}^n)$ by $h(\vec{y}_1^n, \vec{y}_2^n, \vec{y}_3^n | \mathcal{G}^n)$. By rearranging both sides of (B.18), proof of Lemma 11 will be complete.

APPENDIX C

APPENDIX FOR CHAPTER 4

C.1 Proof of Converse for Theorem 6

For each CSIT configuration considered in Table 4.1 we provide the converse proof. Note that the converse proof for the cases *PDD* and *PDN* are already provided in Section 4.3.1. Furthermore, since for the case of *PPP* the only bound is $0 \le d_1, d_2, d_3 \le 1$ according to Table 4.1, the proof is trivial. We now prove the converse for Theorem 6 for the rest of the CSIT configurations.

C.1.1 *PPD*

Note that as mentioned in Remark 20, in order to prove $\frac{d_1}{2} + \frac{d_2}{4} + d_3 \le 1$ for *PDD*, we did not rely on any specific CSIT assumption with respect to Rx_2. Therefore, the bound $\frac{d_1}{2} + \frac{d_2}{4} + d_3 \le 1$ also holds for the case of *PPD*. Moreover, note that by symmetry one can conclude that $\frac{d_1}{4} + \frac{d_2}{2} + d_3 \le 1$ also holds for *PPD*. Hence, since $\frac{d_1}{2} + \frac{d_2}{4} + d_3 \le 1$ and $\frac{d_1}{4} + \frac{d_2}{2} + d_3 \le 1$ constitute the LDoF region for *PPD* according to Table 4.1, the derivations in the converse proof of *PDD* also prove the converse for *PPD*.

C.1.2 *PPN*

According to Table 4.1, it is sufficient to show that $d_1 + d_3 \le 1$ and $d_2 + d_3 \le 1$. We only show $d_1 + d_3 \le 1$; since $d_2 + d_3 \le 1$ can be proven similarly due to symmetry. Suppose (d_1, d_2, d_3) is linearly achievable as defined in Definition 10. Thus, according to (4.10), it is sufficient to show that $m_1(n) + m_3(n) \overset{a.s.}{\leq} n$. By the Decodability condition in (4.9) we have,

$$\begin{aligned}
m_1(n) + m_3(n) &\overset{\substack{(4.9)\\ a.s.}}{=} \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + m_3(n) \\
&\overset{\substack{(4.9)\\ a.s.}}{=} \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + \operatorname{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] - \operatorname{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
&\overset{\text{(Lemma 2)}}{\leq} \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + \operatorname{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_3^n]] - \operatorname{rank}[\mathbf{G}_3^n\mathbf{V}_1^n] \\
&\overset{\substack{\text{(Lemma 16)}\\ a.s.}}{\leq} \operatorname{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_3^n]] \leq n,
\end{aligned}$$

which completes the proof of converse for the case of *PPN*.

C.1.3 *PNN*, *DNN*, *NNN*

According to Table 4.1, it is sufficient to show that $d_1 + d_2 + d_3 \leq 1$. In addition, note that it is sufficient to prove $d_1 + d_2 + d_3 \leq 1$ for the case of *PNN*; since any upper bound for *PNN* is also a valid bound for *DNN* and *NNN*. Suppose (d_1, d_2, d_3) is linearly achievable as defined in Definition 10. Then, according to (4.10), it is sufficient to show that $m_1(n) + m_2(n) + m_3(n) \overset{a.s.}{\leq} n$. By the Decodability condition in (4.9) we have,

$$\begin{aligned}
m_1(n) + m_2(n) + m_3(n) &\overset{a.s.}{=} \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + m_2(n) + m_3(n) \\
&\overset{\substack{(4.9)\\ a.s.}}{=} \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + \operatorname{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \\
&\quad - \operatorname{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_3^n]] + m_3(n) \\
&\overset{\text{(Lemma 2)}}{\leq} \operatorname{rank}[\mathbf{G}_1^n\mathbf{V}_1^n] + \operatorname{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] - \operatorname{rank}[\mathbf{G}_2^n\mathbf{V}_1^n] \\
&\quad + m_3(n) \\
&\overset{\substack{(a)\\ a.s.}}{\leq} \operatorname{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + m_3(n) \\
&\overset{\substack{(4.9)\\ a.s.}}{=} \operatorname{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + \operatorname{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]]
\end{aligned}$$

$$-\mathrm{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]]$$
$$\overset{\substack{(\text{Lemma 16}) \\ a.s.}}{\leq} \mathrm{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \leq n,$$

where (a) follows by applying Lemma 16 and considering Rx_2 as the receiver which supplies no CSIT. Hence, the proof of converse for the cases PNN, DNN, NNN is complete.

C.1.4 *DDD*

The bounds stated in Table 4.1 for *DDD* have been proven in [65] for general encoding schemes via network enhancement and using the fact that in physically degraded broadcast channel feedback does not increase the capacity. Therefore, the same bounds also hold for the class of linear schemes. See [65] for the bounds on the DoF of k-user MISO broadcast channel with delayed CSIT.

C.1.5 *DDN*

Note that according to Table 4.1 and due to symmetry of the first two users it is sufficient to show that $\frac{d_1}{2} + d_2 + d_3 \leq 1$. The other inequality (i.e. $d_1 + \frac{d_2}{2} + d_3 \leq 1$) can be proven similarly. Suppose (d_1, d_2, d_3) is linearly achievable, as defined in Definition 10. Thus, according to (4.10), it is sufficient to show that $\frac{m_1(n)}{2} + m_2(n) + m_3(n) \overset{a.s.}{\leq} n$. We have

$$\begin{aligned}
\frac{m_1(n)}{2} + m_2(n) + m_3(n) \quad & \overset{\substack{(4.9) \\ a.s.}}{=} \quad \frac{\mathrm{rank}[\mathbf{G}_1^n \mathbf{V}_1^n]}{2} + m_2(n) + m_3(n) \\
& \overset{\substack{(4.9) \\ a.s.}}{=} \quad \frac{\mathrm{rank}[\mathbf{G}_1^n \mathbf{V}_1^n]}{2} + \mathrm{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \\
& \qquad -\mathrm{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_3^n]] + m_3(n)
\end{aligned}$$

$$
\begin{aligned}
&\overset{\text{(Lemma 2)}}{\leq} && \frac{\text{rank}[\mathbf{G}_1^n\mathbf{V}_1^n]}{2} + \text{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
& && -\text{rank}[\mathbf{G}_2^n\mathbf{V}_1^n] + m_3(n) \\
&\leq && \frac{\text{rank}[[\mathbf{G}_1^n;\mathbf{G}_2^n]\mathbf{V}_1^n]}{2} + \text{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
& && -\text{rank}[\mathbf{G}_2^n\mathbf{V}_1^n] + m_3(n) \\
&\overset{(a)}{\underset{a.s.}{\leq}} && \text{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + m_3(n) \\
&\overset{(4.9)}{\underset{a.s.}{=}} && \text{rank}[\mathbf{G}_2^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + \text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \\
& && -\text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \\
&\overset{\text{(Lemma 16)}}{\underset{a.s.}{\leq}} && \text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n \quad \mathbf{V}_3^n]] \leq n,
\end{aligned}
$$

where (a) follows by applying Lemma 15 to Rx_2 as the receiver which supplies delayed CSIT. Hence, the proof of converse for the case of *DDN* is complete.

C.2 Proof of Interference Decomposition Bound (Proof of Lemmas 14,17)

Note that Lemma 14 is a special case of Lemma 17 where $k = 3$, $\mathcal{S} = \{1, 2\}$, $j = 3$, and $\ell = 1$. Therefore, in order to prove Lemma 14 and Lemma 17 it is sufficient to prove only Lemma 17. We first restate Lemma 17 here for convenience.

Lemma 17. (Interference Decomposition Bound) *Consider a fixed linear coding strategy $f^{(n)}$, with corresponding precoding matrices $\mathbf{V}_1^n, \mathbf{V}_2^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4). For any $\mathcal{S} \subseteq \{1, 2, \ldots, k\}$, any $\ell \in \mathcal{S}$, and any $j \notin \mathcal{S}$ for which $I_j = D$,*

$$
\frac{\text{rank}[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]] - \text{rank}[\mathbf{G}_\ell^n[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i^n]] + \text{rank}[\mathbf{G}_j^n[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i^n]]}{2} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]]. \tag{C.1}
$$

To prove Lemma 17, we first introduce some definitions. Consider a fixed linear encoding function $f^{(n)}$, with corresponding precoding matrices $\mathbf{V}_1^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4).

Definition 15. *For $\mathcal{S} \subseteq \{1, \ldots, k\}$, $\ell \in \mathcal{S}$, $j \in \{1, \ldots, k\}$, we define*

$$\mathcal{T}_1 \triangleq \{t \in \{1, \ldots, n\} \quad s.t. \quad \mathrm{rank}[\mathbf{G}_\ell^t[\cup_{i\in\mathcal{S}} \mathbf{V}_i^t]] = \mathrm{rank}[\mathbf{G}_\ell^{t-1}[\cup_{i\in\mathcal{S}} \mathbf{V}_i^{t-1}]] + 1\}$$
$$\mathcal{T}_2 \triangleq \{t \in \mathcal{T}_1 \quad s.t. \quad [\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}} \mathbf{V}_i(t)]] \in \mathrm{rowspan}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}} \mathbf{V}_i^{t-1}]]\}.$$

Remark 30. *$\mathcal{T}_1$ is the subset of time slots in which the dimension of received signal at Rx_ℓ increases, while $\mathcal{T}_2$ is the subset of $\mathcal{T}_1$ in which the received signal at Rx_ℓ is already recoverable by using the past received signals at Rx_j. The definitions of $\mathcal{T}_1, \mathcal{T}_2$ focus only on the contribution of $\mathbf{V}_i^n$, where $i \in \mathcal{S}$, on the dimension of received signals at different receivers; because the statement of Lemma 17 only involves $\mathbf{V}_i^n$, where $i \in \mathcal{S}$.*

We now state two lemmas that are the main building blocks of the proof of Lemma 17.

Lemma 28.

$$\mathrm{rank}[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]] - |\mathcal{T}_2| \overset{a.s.}{\leq} \mathrm{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]]. \tag{C.2}$$

Lemma 29.

$$|\mathcal{T}_2| - \mathrm{rank}[\mathbf{G}_\ell^n[\cup_{\substack{i\in\mathcal{S} \\ i\neq\ell}} \mathbf{V}_i^n]] \leq \mathrm{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]] - \mathrm{rank}[\mathbf{G}_j^n[\cup_{\substack{i\in\mathcal{S} \\ i\neq\ell}} \mathbf{V}_i^n]]. \tag{C.3}$$

Note that proof of Lemma 17 is immediate from summing the inequalities in Lemma 28 and Lemma 29. Hence, we will prove Lemma 28 and Lemma 29.

C.2.1 Proof of Lemma 28

Before proving Lemma 28, we first provide its proof sketch for the special case of $k = 3, j = 3, \ell = 1, \mathcal{S} = \{1, 2\}$, the same special case as considered in Lemma 1, to emphasize the underlying ideas. For such special case, Lemma 28 reduces to the following inequality:

$$\text{rank}[\mathbf{G}_1^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] - |\mathcal{T}_2| \overset{a.s.}{\leq} \text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]], \tag{C.4}$$

which can be re-written in the following equivalent form:

$$n - \text{rank}[\mathbf{G}_3^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] \overset{a.s.}{\leq} n - \text{rank}[\mathbf{G}_1^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]] + |\mathcal{T}_2|. \tag{C.5}$$

Note that the L.H.S. of (C.5) is basically the number of time slots $t \in \{1, \ldots, n\}$ in which $\text{rank}[\mathbf{G}_3^t[\mathbf{V}_1^t \quad \mathbf{V}_2^t]]$ does not increase (compared to $\text{rank}[\mathbf{G}_3^{t-1}[\mathbf{V}_1^{t-1} \quad \mathbf{V}_2^{t-1}]]$). Let us denote the set of such time slots by $\mathcal{T}$. First, note that in each $t \in \mathcal{T}$, either $\text{rank}[\mathbf{G}_1^t[\mathbf{V}_1^t \quad \mathbf{V}_2^t]]$ increases by 1 (compared to $\text{rank}[\mathbf{G}_1^{t-1}[\mathbf{V}_1^{t-1} \quad \mathbf{V}_2^{t-1}]]$), or it remains constant. Accordingly, we partition $\mathcal{T}$ into two sets, and upper bound the cardinality of each set. The number of those time slots $t \in \mathcal{T}$ in which $\text{rank}[\mathbf{G}_1^t[\mathbf{V}_2^t \quad \mathbf{V}_2^t]]$ remains constant is at most $n - \text{rank}[\mathbf{G}_1^n[\mathbf{V}_1^n \quad \mathbf{V}_2^n]]$, which constitutes the first two terms on the R.H.S of (C.5).

We now upper bound the number of time slots $t \in \mathcal{T}$, in which $\text{rank}[\mathbf{G}_1^t[\mathbf{V}_2^t \quad \mathbf{V}_2^t]]$ increases by 1. In each such time slot, Rx_3 receives an equation which is already recoverable by using its past received equations (since $\text{rank}[\mathbf{G}_3^t[\mathbf{V}_1^t \quad \mathbf{V}_2^t]]$ does not increase). But note that due to the assumption of delayed CSIT for Rx_3, the transmitter does not know the channels to Rx_3 when transmitting its signals at time slot t; and the received signal at Rx_3 would be a random linear combination of transmit signals. In order for this random linear combination to be known at Rx_3, Rx_3 must have already been able to recover each of the individual signals transmitted at time t, based on its past received signals. Note that if Rx_3 knows each individual transmit signal at time t, it also knows any linear combination of them. Hence, it can already recover what Rx_1 receives at time t. Therefore, the number of time slots $t \in \mathcal{T}$ in which $\text{rank}[\mathbf{G}_1^t[\mathbf{V}_2^t \quad \mathbf{V}_2^t]]$ increases by 1 is upper bounded by the number of time slots in which the received signal at Rx_1 is already recoverable by Rx_3, and $\text{rank}[\mathbf{G}_1^t[\mathbf{V}_2^t \quad \mathbf{V}_2^t]]$ increases by 1, which in turn, by the definition of $\boldsymbol{\mathcal{T}_2}$ is equal to $|\boldsymbol{\mathcal{T}_2}|$, the last term on the R.H.S. of (C.5). Thus, the proof sketch is complete.

The following is the general mathematical proof for Lemma 28, which relies on the above approach. Let us denote the indicator function by $I(.)$. We then have

$$
\begin{aligned}
n - \text{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]] &= \sum_{t=1}^{n} I(\text{rank}[\mathbf{G}_j^t[\cup_{i\in\mathcal{S}}\mathbf{V}_i^t]] = \text{rank}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&= \sum_{t\in\boldsymbol{\mathcal{T}_1}} I(\text{rank}[\mathbf{G}_j^t[\cup_{i\in\mathcal{S}}\mathbf{V}_i^t]] = \text{rank}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&\quad + \sum_{t\in\boldsymbol{\mathcal{T}_1^c}} I(\text{rank}[\mathbf{G}_j^t[\cup_{i\in\mathcal{S}}\mathbf{V}_i^t]] = \text{rank}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&\overset{\substack{(a)\\ a.s.}}{=} \sum_{t\in\boldsymbol{\mathcal{T}_1}} I(\text{rowspan}[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)] \subseteq \text{rowspan}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&\quad + \sum_{t\in\boldsymbol{\mathcal{T}_1^c}} I(\text{rank}[\mathbf{G}_j^t[\cup_{i\in\mathcal{S}}\mathbf{V}_i^t]] = \text{rank}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&\leq \sum_{t\in\boldsymbol{\mathcal{T}_1}} I(\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)] \in \text{rowspan}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&\quad + \sum_{t\in\boldsymbol{\mathcal{T}_1^c}} I(\text{rank}[\mathbf{G}_j^t[\cup_{i\in\mathcal{S}}\mathbf{V}_i^t]] = \text{rank}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&= |\boldsymbol{\mathcal{T}_2}| + \sum_{t\in\boldsymbol{\mathcal{T}_1^c}} I(\text{rank}[\mathbf{G}_j^t[\cup_{i\in\mathcal{S}}\mathbf{V}_i^t]] = \text{rank}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]) \\
&\leq |\boldsymbol{\mathcal{T}_2}| + \sum_{t\in\boldsymbol{\mathcal{T}_1^c}} 1 = |\boldsymbol{\mathcal{T}_2}| + n - |\boldsymbol{\mathcal{T}_1}|
\end{aligned}
$$

$$\overset{(b)}{=} |\mathcal{T}_{\mathbf{2}}| + n - \text{rank}[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]],$$

where (a) is due to Lemma 30, which is stated and proved in Appendix C.3 [1]; and (b) follows immediately from the definition of $\mathcal{T}_{\mathbf{1}}$. By rearranging the above inequality, the proof of Lemma 28 will be complete.

C.2.2 Proof of Lemma 29

We first state a claim which is useful in lower bounding the R.H.S. of the inequality in Lemma 29, and it can be proved using simple linear algebra; hence the proof is omitted for brevity.

Claim 6. *For two matrices A, B of the same row size,* $\text{rank}[A \quad B] - \text{rank}[B] = \dim(\text{span}([\vec{s} \quad \vec{0}]$ *s.t.* $[\vec{s} \quad \vec{0}] \in \text{rowspan}[A \quad B]))$.

We are now ready to prove Lemma 29. Let $[\vec{\mathbf{g}}_\ell(t)[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_{\mathbf{2}}}$ denote the matrix constructed by rows $\vec{\mathbf{g}}_\ell(t)[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i(t)]$, where $t \in \mathcal{T}_{\mathbf{2}}$. We have

$$\begin{aligned}
& \text{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]]] - \text{rank}[\mathbf{G}_j^n[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i^n]]] \\
\overset{\text{(Claim 6)}}{=} \quad & \dim(\text{span}([\vec{s} \quad \vec{0}] \text{ s.t. } [\vec{s} \quad \vec{0}] \in \text{rowspan}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]])) \\
\overset{(a)}{\geq} \quad & \dim(\text{span}([\vec{s} \quad \vec{0}] \text{ s.t. } [\vec{s} \quad \vec{0}] \in \text{rowspan}[[\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_{\mathbf{2}}}])) \\
\overset{\text{(Claim 6)}}{=} \quad & \text{rank}[[\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_{\mathbf{2}}}] - \text{rank}[[\vec{\mathbf{g}}_\ell(t)[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_{\mathbf{2}}}] \\
\overset{(b)}{=} \quad & |\mathcal{T}_{\mathbf{2}}| - \text{rank}[[\vec{\mathbf{g}}_\ell(t)[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_{\mathbf{2}}}] \\
\geq \quad & |\mathcal{T}_{\mathbf{2}}| - \text{rank}[\mathbf{G}_\ell^n[\cup_{\substack{i\in\mathcal{S}\\ i\neq\ell}}\mathbf{V}_i^n]],
\end{aligned}$$

[1] Lemma 30 is a variation of Lemma 25 in Appendix A.2 which was stated for the setting with *distributed* transmit antennas. Proof of Lemma 30 follows similar steps as in the proof of Lemma 25; it is provided in Appendix C.3 for completeness.

where (a) follows from the fact that for each $t \in \mathcal{T}_2$, $\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)] \in \text{rowspan}[\mathbf{G}_j^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]]$; hence, for each $t \in \mathcal{T}_2$, $\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)] \in \text{rowspan}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]]$; and therefore,

$$\text{rowspan}[[\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_2}] \subseteq \text{rowspan}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]].$$

Furthermore, (b) holds since $\text{rank}[[\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_2}] = |\mathcal{T}_2|$, which is due to the following: since $\mathcal{T}_2 \subseteq \mathcal{T}_1$, if $t \in \mathcal{T}_2$, then $t \in \mathcal{T}_1$. Therefore, using the definition of $\mathcal{T}_1$, we get

$$\forall t \in \mathcal{T}_2, \qquad \vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)] \notin \text{rowspan}\left(\mathbf{G}_\ell^{t-1}[\cup_{i\in\mathcal{S}}\mathbf{V}_i^{t-1}]\right). \tag{C.6}$$

Consequently, the vectors $\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)]$, where $t \in \mathcal{T}_2$, are linearly independent; and therefore, we have $\text{rank}[[\vec{\mathbf{g}}_\ell(t)[\cup_{i\in\mathcal{S}}\mathbf{V}_i(t)]]_{t\in\mathcal{T}_2}] = |\mathcal{T}_2|$. Hence, the proof of Lemma 29 is complete.

C.3 Statement and Proof of Lemma 30

Lemma 30. *Consider a fixed linear coding strategy $f^{(n)}$ with corresponding precoding matrices $\mathbf{V}_1^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4). Consider an arbitrary index j, where $j \in \{1, \ldots, k\}$, and assume $I_j \in \{D, N\}$; i.e., the transmitter has either delayed or no CSIT with respect to Rx$_j$. In addition, consider an arbitrary set of receiver indices $\mathcal{S}$, where $\mathcal{S} \subseteq \{1, \ldots, k\}$. For any $t \in \{1, 2, \ldots, n\}$, let $\mathcal{A}_t, \mathcal{B}_t$, denote the following sets of channel realizations:*

- $\mathcal{A}_t \triangleq \{\mathcal{G}^n | \quad \text{rank}[G_j^t[\cup_{i\in\mathcal{S}}V_i^t]] = \text{rank}[G_j^{t-1}[\cup_{i\in\mathcal{S}}V_i^{t-1}]]\}$.
- $\mathcal{B}_t \triangleq \{\mathcal{G}^n | \quad \text{rowspan}[\cup_{i\in\mathcal{S}}V_i(t)] \subseteq \text{rowspan}[G_j^{t-1}[\cup_{i\in\mathcal{S}}V_i^{t-1}]]\}$.

Then,

$$\Pr(\boldsymbol{\mathcal{G}}^n \in \cup_{t=1}^n(\mathcal{A}_t \cap \mathcal{B}_t^c)) = 0.$$

Proof. Note that due to Union Bound, it is sufficient to show that for any $t \in \{1, 2, \ldots, n\}$, $\Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_t \cap \mathcal{B}_t^c) = 0$. Consider an arbitrary $t \in \{1, 2, \ldots, n\}$. Due to Total Probability Law, it is sufficient to show that for any channel realization of the first $t-1$ timeslots, denoted by $\mathcal{G}^{t-1}$, we have

$$\Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_t \cap \mathcal{B}_t^c | \boldsymbol{\mathcal{G}}^{t-1} = \mathcal{G}^{t-1}) = 0. \tag{C.7}$$

Consider an arbitrary channel realization of the first $t-1$ time slots $\mathcal{G}^{t-1}$ and precoding matrices $V_1^t, \ldots, V_k^t$ (which are now deterministic because they are only function of the channel realizations for the first $t-1$ time slots). Also, suppose that given $\mathcal{G}^{t-1}$, $\mathcal{B}_t^c$ occurs; since otherwise, the proof of (C.7) would be complete. We denote the row h of the matrix $[\cup_{i\in\mathcal{S}} V_i(t)]$ by $[\cup_{i\in\mathcal{S}} V_{i,h}(t)]$. Note that by assuming $\mathcal{B}_t^c$ occurs, and denoting $\mathcal{L} = \text{rowspan}[G_j^{t-1}[\cup_{i\in\mathcal{S}} V_i^{t-1}]]$, the following is true (according to the definition of $\mathcal{B}_t$):

$$\begin{aligned} & \exists h \in \{1, \ldots, m\} \quad s.t. \quad [\cup_{i\in\mathcal{S}} V_{i,h}(t)] \notin \mathcal{L} \\ \Rightarrow \quad & \exists h \in \{1, \ldots, m\} \quad s.t. \quad \text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,h}(t)] \neq 0. \end{aligned} \tag{C.8}$$

Therefore, the $m \times (\sum_{i\in\mathcal{S}} m_i(n))$ matrix $[\text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,1}(t)]; \ldots; \text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,m}(t)]]$ is non-zero, which means that its null space has dimension strictly lower than m, the number of its rows. Hence, we have,

$$\begin{aligned} \Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_t \cap \mathcal{B}_t^c | \boldsymbol{\mathcal{G}}^{t-1} = \mathcal{G}^{t-1}) &\overset{(a)}{=} \Pr(\boldsymbol{\mathcal{G}}^n \in \mathcal{A}_t | \boldsymbol{\mathcal{G}}^{t-1} = \mathcal{G}^{t-1}) \\ &\overset{(b)}{=} \Pr(\text{Proj}_{\mathcal{L}^\perp}[\vec{\mathbf{g}}_j(t)[\cup_{i\in\mathcal{S}} V_i(t)]] = 0 | \boldsymbol{\mathcal{G}}^{t-1} = \mathcal{G}^{t-1}) \\ &\overset{(c)}{=} \Pr(\vec{\mathbf{g}}_j(t)[\text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,1}(t)]; \ldots; \text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,m}(t)]] = 0 | \boldsymbol{\mathcal{G}}^{t-1} = \mathcal{G}^{t-1}) \\ &= \Pr(\vec{\mathbf{g}}_j(t)^\top \in \text{nullspace}([\text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,1}(t)]; \ldots; \text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}} V_{i,m}(t)]]^\top) \end{aligned}$$

$$|\mathcal{G}^{t-1} = \mathcal{G}^{t-1})$$

$$\stackrel{(d)}{=} 0,$$

where (a) holds since we assumed that for realization $\mathcal{G}^{t-1}$, $\mathcal{B}_t^c$ occurs; (b) holds according to the definition of $\mathcal{A}_t$; (c) holds due to linearity of orthogonal projection; and (d) holds since, as mentioned before, the matrix $[\text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}}V_{i,1}(t)];\ldots;\text{Proj}_{\mathcal{L}^\perp}[\cup_{i\in\mathcal{S}}V_{i,m}(t)]]^\top$ is non-zero, meaning that its null space, which is a subspace in $\mathbb{R}^m$, has dimension strictly lower than m. Therefore, the probability that the random vector $\vec{\mathbf{g}}_j(t)$ lies in a subspace in $\mathbb{R}^m$ of strictly lower dimension (than m) is zero.

□

C.4 Proof of MIMO Rank Ratio Inequality for BC (Proof of Lemmas 15,18)

Note that Lemma 15 is a special case of Lemma 18 where $k = 3$, $j = 1$, $\mathcal{S} = \{i\}$, and $i_1 = 3, i_2 = \ell$. Therefore, in order to prove Lemma 15 and Lemma 18 it is sufficient to prove only Lemma 18. We first re-state Lemma 18 here for convenience.

Lemma 18. (MIMO Rank Ratio Inequality for BC) *Consider a linear coding strategy $f^{(n)}$, with corresponding $\mathbf{V}_1^n,\ldots,\mathbf{V}_k^n$ as defined in (4.4). Let $\mathbf{Y}_j^n \triangleq \mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]$, where $\mathcal{S} \subseteq \{1,2,\ldots,k\}$. Also, consider distinct receivers $Rx_{i_1},\ldots,Rx_{i_{j+1}}$, where $j = 1,2,\ldots,k-1$ and $i_1,\ldots,i_{j+1} \in \{1,\ldots,k\}$. If $Rx_{i_1},\ldots,Rx_{i_j}$ supply delayed CSIT, then,*

$$\frac{\text{rank}[\mathbf{Y}_{i_1}^n;\ldots;\mathbf{Y}_{i_{j+1}}^n]}{j+1} \stackrel{a.s.}{\leq} \frac{\text{rank}[\mathbf{Y}_{i_1}^n;\ldots;\mathbf{Y}_{i_j}^n]}{j}. \tag{C.9}$$

Proof. Without loss of generality, we suppose that $i_1 = 1, i_2 = 2, \dots, i_{j+1} = j+1$. Thus, we need to show that

$$\frac{\text{rank}[\mathbf{Y}_1^n; \dots; \mathbf{Y}_{j+1}^n]}{j+1} \overset{a.s.}{\leq} \frac{\text{rank}[\mathbf{Y}_1^n; \dots; \mathbf{Y}_j^n]}{j}. \tag{C.10}$$

Let us denote

$$\text{rank}[A|B] \triangleq \text{rank}[A; B] - \text{rank}[B]. \tag{C.11}$$

Hence, by sub-modularity property of rank (Lemma 2), for matrices A, B, C with the same number of columns,

$$\text{rank}[A|B] \geq \text{rank}[A|B; C]; \tag{C.12}$$

$$\text{rank}[A|C] + \text{rank}[B|C] \geq \text{rank}[A; B|C]. \tag{C.13}$$

Moreover, we denote $\mathbf{Y}_j(t) \triangleq \vec{\mathbf{g}}_j(t)[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)]$ and $\mathbf{Y}^t \triangleq [\mathbf{Y}_1^t; \dots; \mathbf{Y}_j^t]$. For each $i = 1, \dots, j$, we have

$$\begin{aligned}
&\text{rank}[\mathbf{Y}_i(t)|[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]] \\
\overset{(C.11)}{=}\; & I\left(\vec{\mathbf{g}}_i(t)[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \notin \text{rowspan}[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]\right) && \text{(C.14)} \\
=\; & 1 - I\left(\vec{\mathbf{g}}_i(t)[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \in \text{rowspan}[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]\right) && \text{(C.15)} \\
\overset{(a)}{\underset{a.s.}{=}}\; & 1 - I\left(\text{rowspan}[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \subseteq \text{rowspan}[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]\right) && \text{(C.16)} \\
\overset{(b)}{\geq}\; & 1 - I\left(\vec{\mathbf{g}}_{j+1}(t)[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \in \text{rowspan}[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]\right) && \text{(C.17)} \\
\overset{(C.11)}{=}\; & \text{rank}[\mathbf{Y}_{j+1}(t)|[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]] && \text{(C.18)} \\
\overset{(C.12)}{\geq}\; & \text{rank}[\mathbf{Y}_{j+1}(t)|[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_j(t)]] && \text{(C.19)} \\
\overset{(C.11)}{=}\; & \text{rank}[[\mathbf{Y}_1(t); \dots; \mathbf{Y}_{j+1}(t)]|\mathbf{Y}^{t-1}] - \text{rank}[[\mathbf{Y}_1(t); \dots; \mathbf{Y}_j(t)]|\mathbf{Y}^{t-1}] && \text{(C.20)} \\
\overset{(C.12)}{\geq}\; & \text{rank}[[\mathbf{Y}_1(t); \dots; \mathbf{Y}_{j+1}(t)]|[\mathbf{Y}^{t-1}; \mathbf{Y}_{j+1}^{t-1}]] - \text{rank}[[\mathbf{Y}_1(t); \dots; \mathbf{Y}_j(t)]|\mathbf{Y}^{t-1}], \\
& && \text{(C.21)}
\end{aligned}$$

where to see why (a) holds, we first present the following variant of Lemma 30: if $\mathcal{A}_t$ denotes the event: $\vec{\mathbf{g}}_i(t)[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \in \text{rowspan}[\mathbf{Y}^{t-1}; \mathbf{Y}_1(t); \dots; \mathbf{Y}_{i-1}(t)]$, and $\mathcal{B}_t$

denotes the event: $\text{rowspan}[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \subseteq \text{rowspan}[\mathbf{Y}^{t-1};\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{i-1}(t)]$, then

$$\Pr\left(\mathcal{A}_t \cap \mathcal{B}_t^c\right) = 0, \tag{C.22}$$

which can be proven using the same steps as in proof of Lemma 30; therefore, its proof is omitted for brevity. As a result, $I(\mathcal{A}_t) = I(\mathcal{A}_t \cap \mathcal{B}_t) + I(\mathcal{A}_t \cap \mathcal{B}_t^c) \overset{(C.22)}{\underset{a.s.}{=}} I(\mathcal{A}_t \cap \mathcal{B}_t) = I(\mathcal{B}_t)$, where the last equality holds since occurrence of $\mathcal{B}_t$ implies occurrence of $\mathcal{A}_t$. Therefore, $1 - I(\mathcal{A}_t) \overset{a.s.}{=} 1 - I(\mathcal{B}_t)$. In addition, note that the left-hand-side of (a) is $1 - I(\mathcal{A}_t)$, and the right-hand-side of (a) is $1 - I(\mathcal{B}_t)$. Hence, (a) holds. Moreover, (b) holds due to the fact that if

$$\text{rowspan}[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \subseteq \text{rowspan}[\mathbf{Y}^{t-1};\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{i-1}(t)],$$

then,

$$\vec{\mathbf{g}}_{j+1}(t)[\cup_{h\in\mathcal{S}}\mathbf{V}_h(t)] \in \text{rowspan}[\mathbf{Y}^{t-1};\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{i-1}(t)].$$

By summing both sides of (C.21) over $i = 1,\ldots,j$, and using (C.11) we obtain,

$$\begin{aligned}&\text{rank}[[\mathbf{Y}_1(t);\ldots;\mathbf{Y}_j(t)]|\mathbf{Y}^{t-1}]\\ &\overset{a.s.}{\geq} j\left(\text{rank}[[\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{j+1}(t)]|[\mathbf{Y}^{t-1};\mathbf{Y}_{j+1}^{t-1}]] - \text{rank}[[\mathbf{Y}_1(t);\ldots;\mathbf{Y}_j(t)]|\mathbf{Y}^{t-1}]\right);\end{aligned}$$

and by rearranging the above inequality and dividing both sides by $j(j+1)$ we obtain

$$\frac{\text{rank}[[\mathbf{Y}_1(t);\ldots;\mathbf{Y}_j(t)]|\mathbf{Y}^{t-1}]}{j} \overset{a.s.}{\geq} \frac{\text{rank}[[\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{j+1}(t)]|[\mathbf{Y}^{t-1};\mathbf{Y}_{j+1}^{t-1}]]}{j+1}. \tag{C.23}$$

Finally, by summing both sides of the above inequality over all $t = 1,\ldots,n$, and using the definition in (C.11), the proof of (C.10) would be complete, which concludes the proof of Lemma 18. □

C.5 Proof of Least Alignment Lemma (Proof of Lemmas 16,19)

Note that Lemma 16 is a special case of Lemma 19 where $k = 3$ and $j = 3$. Therefore, in order to prove Lemma 16 and Lemma 19 it is sufficient to prove only Lemma 19. We first re-state Lemma 19 here for convenience.

Lemma 19. **(Least Alignment Lemma)** *For any linear coding strategy $f^{(n)}$, with corresponding $\mathbf{V}_1^n, \ldots, \mathbf{V}_k^n$ as defined in (4.4), and any $\mathcal{S} \subseteq \{1, 2, \ldots, k\}$, if $I_j = N$ for some $j \in \{1, 2, \ldots, k\}$,*

$$\forall \ell \in \{1, 2, \ldots, k\}, \qquad \operatorname{rank}\left[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]\right] \overset{a.s.}{\leq} \operatorname{rank}\left[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]\right].$$

Proof. Define $m(n) \triangleq \sum_{i\in\mathcal{S}} m_i(n)$. We first state a lemma that will be later useful in proving Lemma 19.

Lemma 31. **([22])** *For $n \in \mathbb{N}$, a multi-variate polynomial function on $\mathbb{C}^n$ to $\mathbb{C}$, is either identically* 0*, or non-zero almost everywhere.*

We now prove Lemma 19. Denote by $[1 : n]$ the set $\{1, \ldots, n\}$. For any matrix $B_{n\times m(n)}$ and $I_1 \subseteq [1 : n]$, and $I_2 \subseteq [1 : m(n)]$, we denote by B_{I_1,I_2} the sub-matrix of B whose rows and columns are specified by I_1 and I_2, respectively. Define the set of channel realizations $\mathcal{A}$ as:

$$\mathcal{A} \triangleq \left\{\mathcal{G}^n | \operatorname{rank}[G_\ell^n[\cup_{i\in\mathcal{S}} V_i^n]] > \operatorname{rank}[G_j^n[\cup_{i\in\mathcal{S}} V_i^n]]\right\}. \tag{C.24}$$

In order to prove $\operatorname{rank}[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]] \overset{a.s.}{\leq} \operatorname{rank}[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}} \mathbf{V}_i^n]]$, we only need to show $\Pr(\mathcal{A}) = 0$. Since a matrix $B_{n\times m(n)}$ has rank r if and only if the maximum size of a square sub-matrix of B with non-zero determinant is r,

$$\begin{aligned}\mathcal{A} \subseteq \{\mathcal{G}^n| \quad & \exists I_1 \subseteq [1:n], I_2 \subseteq [1:m(n)], |I_1| = |I_2|,\\ & s.t. \quad \det([G_\ell^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \neq \det([G_j^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) = 0\},\end{aligned}$$

which can be rewritten as

$$A \subseteq \cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m(n)]\\ |I_1|=|I_2|}} \left\{\mathcal{G}^n|\det([G_\ell^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \neq 0, \quad \det([G_j^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) = 0\right\}. \tag{C.25}$$

Let X^n denote a diagonal matrix of size $n\times n$ where the elements on the diagonal are variables in $\mathbb{C}$. Then, for any $I_1 \subseteq [1:n], I_2 \subseteq [1:m(n)],$ where $|I_1| = |I_2|$, $\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2})$ is a multi-variate polynomial function in the elements of X^n. Note that if for some realization $X^n = G_\ell^n$, $\det([G_\ell^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \neq 0$, then the polynomial function defined by $\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2})$ is not identical to zero (i.e., $\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \overset{\text{identical}}{\neq} 0$). So, by (C.25), we have

$$\begin{aligned}\mathcal{A} &\subseteq \cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m(n)]\\ |I_1|=|I_2|}} \{\mathcal{G}^n|\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \overset{\text{identical}}{\neq} 0, \quad \det([G_j^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) = 0\}\\ &= \cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m(n)]\\ |I_1|=|I_2|}} \{\mathcal{G}^n|\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \overset{\text{identical}}{\neq} 0, \quad G_j^n \text{ is root of } \det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2})\}.\end{aligned} \tag{C.26}$$

Note that by Lemma 31, for every $I_1 \in [1:n], I_2 \in [1:m(n)], |I_1| = |I_2|$, we have

$$\Pr(\{\mathcal{G}^n|\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \overset{\text{identical}}{\neq} 0, \quad G_j^n \text{ is root of } \det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2})\}) = 0. \tag{C.27}$$

So, since finite union of measure-zero sets has measure zero,

$$\Pr(\cup_{\substack{I_1\subseteq[1:n]\\ I_2\subseteq[1:m(n)]\\ |I_1|=|I_2|}} \{\mathcal{G}^n|\det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2}) \overset{\text{identical}}{\neq} 0, G_j^n\text{: root of } \det([X^n[\cup_{i\in\mathcal{S}} V_i^n]]_{I_1,I_2})\}) = 0, \tag{C.28}$$

which by (C.26) implies that $\Pr(\mathcal{A}) = 0$. Therefore, according to the definition of $\mathcal{A}$ in (C.24),

$$\operatorname{rank}\left[\mathbf{G}_\ell^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]\right] \overset{a.s.}{\leq} \operatorname{rank}\left[\mathbf{G}_j^n[\cup_{i\in\mathcal{S}}\mathbf{V}_i^n]\right], \tag{C.29}$$

which completes the proof of Least Alignment Lemma. □

Remark 31. *Using the same line of argument as in the proof of Lemma 19, one can prove Lemma 19 for a more general network setting where there are arbitrary number of transmitters, and the transmitters have arbitrary number of antennas. In addition, the statement of Lemma 19 holds even if Rx_j, Rx_ℓ have multiple but equal number of antennas.*

C.6 Proof of Proposition 6 (Constant Gap Characterization for $|\mathcal{P}| \geq |\mathcal{D}|$)

In this Appendix we show that for $|\mathcal{P}| \geq |\mathcal{D}|$, Theorem 7 leads to an approximate characterization of LDoF_{sum} to within an additive gap of $\frac{1}{2}$, as presented in Proposition 6. First, note that for the special case of $|\mathcal{P}| = |\mathcal{D}| = 0$, $\text{LDoF}_{\text{region}}$ is completely characterized by $\{(d_1, \ldots, d_k) \quad | \quad \sum_{i=1}^{k} d_i \leq 1\}$. Thus, henceforth we assume that $|\mathcal{P}| > 0$.

Moreover, note that a naive lower bound for LDoF_{sum} is $|\mathcal{P}|$; since we can focus only on the $|\mathcal{P}|$ receivers that provide instantaneous CSIT, and for those $|\mathcal{P}|$ receivers we can perform zero-forcing to cancel interference and achieve $|\mathcal{P}|$ as a lower bound on LDoF_{sum}. Using this lower bound we show that for the case where $|\mathcal{P}| \geq |\mathcal{D}|$, the statement of Proposition 6 holds. In particular, we first consider the case where $|\mathcal{D}| = 0$. For this case, by (4.27) in Theorem 7 we have

$$\forall i \in \mathcal{P} \cup \mathcal{D}, \quad d_i + \sum_{j \in \mathcal{N}} d_j \leq 1, \tag{C.30}$$

which, together with $\forall i, d_i \leq 1$, yields

$$\mathrm{LDoF}_{\mathrm{sum}} \leq |\mathcal{P}|. \tag{C.31}$$

Hence, the naive lower bound of $|\mathcal{P}|$ on $\mathrm{LDoF}_{\mathrm{sum}}$ is tight for the case where $|\mathcal{D}| = 0$. Moreover, $\mathrm{LDoF}_{\mathrm{sum}}$ for the special case where $|\mathcal{D}| = 1$ is characterized in Proposition 7. Therefore, we only need to prove Proposition 6 for the case of $|\mathcal{P}| \geq |\mathcal{D}| > 1$. Recall that by (4.25) in Theorem 7,

$$\forall i \in \mathcal{D}, \forall \pi_{\mathcal{P}\cup\mathcal{D}\setminus i}, \qquad \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{d_{\pi_{\mathcal{P}\cup\mathcal{D}\setminus i}(j)}}{2^j} + d_i + \sum_{j \in \mathcal{N}} d_j \leq 1. \tag{C.32}$$

Without loss of generality, suppose $\mathcal{P} = \{1, \ldots, |\mathcal{P}|\}$, and $\mathcal{D} = \{|\mathcal{P}| + 1, \ldots, |\mathcal{P}| + |\mathcal{D}|\}$, and $\mathcal{N} = \{|\mathcal{P}| + |\mathcal{D}| + 1, \ldots, k\}$. In addition, let $i = |\mathcal{P}| + |\mathcal{D}|$, and $\pi_{\mathcal{P}\cup\mathcal{D}\setminus i}$ be the identity permutation. Consequently, by (4.25) in Theorem 7 we obtain:

$$\sum_{i=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{d_i}{2^i} + d_{|\mathcal{P}|+|\mathcal{D}|} + \sum_{j \in \mathcal{N}} d_j \leq 1, \tag{C.33}$$

or equivalently,

$$(\sum_{i=1}^{|\mathcal{P}|} \frac{d_i}{2^i}) + (\sum_{i=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{d_i}{2^i} + d_{|\mathcal{P}|+|\mathcal{D}|}) + \sum_{j \in \mathcal{N}} d_j \leq 1. \tag{C.34}$$

Note that in the above inequality there are $|\mathcal{P}|$ different coefficients (i.e. $\frac{1}{2}, \ldots, \frac{1}{2^{|\mathcal{P}|}}$) for receivers in $\mathcal{P}$, and $|\mathcal{D}|$ different coefficients (i.e. $\frac{1}{2^{|\mathcal{P}|+1}}, \ldots, \frac{1}{2^{|\mathcal{P}|+|\mathcal{D}|-1}}, 1$) for receivers in $\mathcal{D}$. Due to symmetry, we can consider all the possible $|\mathcal{P}|! \times |\mathcal{D}|!$ joint permutations of the receivers in $\mathcal{P}$ and $\mathcal{D}$, leading to permutations of the corresponding coefficients in (C.34). By summing over all those resulting inequalities, and diving by $|\mathcal{P}|! \times |\mathcal{D}|!$, we obtain

$$(1 - \frac{1}{2^{|\mathcal{P}|}})(\sum_{i=1}^{|\mathcal{P}|} \frac{d_i}{|\mathcal{P}|}) + (1 + \frac{1}{2^{|\mathcal{P}|}} - \frac{1}{2^{|\mathcal{P}|+|\mathcal{D}|-1}})(\sum_{i=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|} \frac{d_i}{|\mathcal{D}|}) + \sum_{j \in \mathcal{N}} d_j \leq 1. \tag{C.35}$$

Note that $\text{LDoF}_{\text{sum}} \leq \max \sum_{i=1}^{k} d_i$ subject to (C.35) and $d_i \leq 1$ for all i, which is basically a simple linear program. By solving the linear program, one can easily see that

$$\text{LDoF}_{\text{sum}} \leq \max\left(|\mathcal{P}| + \frac{|\mathcal{D}|}{2^{|\mathcal{P}|} + 1 - \frac{1}{2^{|\mathcal{D}|-1}}}, |\mathcal{P}| + \frac{1}{2^{|\mathcal{P}|}}, 1, \frac{|\mathcal{D}|}{1 + \frac{1}{2^{|\mathcal{P}|}} - \frac{1}{2^{|\mathcal{P}|+|\mathcal{D}|-1}}}\right). \tag{C.36}$$

Note that since we assumed $|\mathcal{P}| \geq |\mathcal{D}| > 1$, the above inequality simplifies as follows:

$$\text{LDoF}_{\text{sum}} \leq |\mathcal{P}| + \frac{|\mathcal{D}|}{2^{|\mathcal{P}|} + 1 - \frac{1}{2^{|\mathcal{D}|-1}}}, \tag{C.37}$$

which together with $\text{LDoF}_{\text{sum}} \geq |\mathcal{P}|$ leads to

$$|\mathcal{P}| \leq \text{LDoF}_{\text{sum}} \leq |\mathcal{P}| + \frac{|\mathcal{D}|}{2^{|\mathcal{P}|} + 1 - \frac{1}{2^{|\mathcal{D}|-1}}}. \tag{C.38}$$

Therefore, the gap between upper and lower bounds on LDoF_{sum} is upper bounded as

$$\text{Gap} = \frac{|\mathcal{D}|}{2^{|\mathcal{P}|} + 1 - \frac{1}{2^{|\mathcal{D}|-1}}} \leq \frac{|\mathcal{D}|}{2^{|\mathcal{P}|}} \leq \frac{|\mathcal{P}|}{2^{|\mathcal{P}|}} \leq \frac{1}{2}.$$

Hence, the proof of Proposition 6 is complete.

C.7 Proof of Proposition 7 ($\text{LDoF}_{\text{sum}} = |\mathcal{P}| + \frac{1}{2^{|\mathcal{P}|}}$ for $|\mathcal{D}| = 1$)

We focus on the k-user MISO BC with only one receiver supplying delayed CSIT. We first prove the converse. Assume without loss of generality that $\mathcal{P} = \{1, \ldots, |\mathcal{P}|\}$, $\mathcal{D} = \{|\mathcal{P}| + 1\}$ and $\mathcal{N} = \{|\mathcal{P}| + 2, \ldots, k\}$. Further, let $i = |\mathcal{P}| + 1$, and $\pi_{\mathcal{P}\cup\mathcal{D}\setminus i}$ denote the identity permutation. Then, by Theorem 7 the solution to the following linear program provides an upper bound on LDoF_{sum}:

$$\text{LDoF}_{\text{sum}} \leq \quad \max \quad \sum_{i=1}^{k} d_i$$

$$s.t. \quad \sum_{i=1}^{|\mathcal{P}|} \frac{d_i}{2^i} + d_{|\mathcal{P}|+1} + \sum_{j \in \mathcal{N}} d_j \leq 1, \tag{C.39}$$

$$0 \leq d_i \leq 1, \qquad i = 1, \ldots, k, \tag{C.40}$$

where the first constraint in the linear program is due to (4.25) in Theorem 7. Thus, by solving the above linear program one can readily see that

$$\mathrm{LDoF}_{\mathrm{sum}} \leq |\mathcal{P}| + \frac{1}{2^{|\mathcal{P}|}}. \tag{C.41}$$

Hence, the converse proof is complete. We now present the achievable scheme, which is a multi-phase scheme that uses hybrid CSIT available to the transmitter to perform interference alignment. The new achievable scheme generalizes the schemes for *PD* in [82] and *PPD* in [9] (see Figure C.1 for the special case of *PPPD*).

To achieve $\mathrm{LDoF}_{\mathrm{sum}}$ of $|\mathcal{P}| + \frac{1}{2^{|\mathcal{P}|}}$, we will ignore the receivers in $\mathcal{N}$; and we show that we can linearly achieve $(d_1, \ldots, d_{|\mathcal{P}|+1}) = \left(1, \ldots, 1, \frac{1}{2^{|\mathcal{P}|}}\right)$. Therefore, if, with slight abuse of notation, we denote $K \triangleq |\mathcal{P}| + 1$, we need to show that the following DoF tuple is linearly achievable:

$$(d_1, \ldots, d_{K-1}, d_K) = \left(1, \ldots, 1, \frac{1}{2^{K-1}}\right). \tag{C.42}$$

To this end, we present a new multi-phase communication scheme which

- operates over 2^{K-1} time slots;
- delivers 2^{K-1} symbols to each of the receivers $1, \ldots, K-1$;
- delivers 1 symbol to receiver K.

The overall scheme is split into K phases, indexed as $i = 0, 1, 2, \ldots, (K-1)$:

- the duration of i-th phase is $\binom{K-1}{i}$ time slots;

- each of the first $(K-1)$ receivers obtain new (interference-free) linear equations in every time slot;
- receiver K obtains $\binom{K-1}{i}$ equations during phase i (one corresponding to each time slot).

At the end of the i-th phase, receiver K does the following: it uses its received $\binom{K-1}{i-1}$ equations from phase $(i-1)$ and $\binom{K-1}{i}$ equations in phase i to obtain $\binom{K-1}{i}$ new equations with the following specific property: each equation is a linear combination of the desired symbol by Rx_K and $(K-1-i)$ undesired symbols, where each undesired symbol is in fact desired by another receiver.

Throughout the proof of the achievable scheme we only utilize the first K transmit antennas; therefore, without loss of generality we can assume as well that there are only K transmit antennas. We first start with Phase 0, and then explain the transmission strategy for an arbitrary phase i in full detail.

C.7.1 Phase 0

Phase 0 is of duration $\binom{K-1}{0}=1$, i.e., this phase only has 1 time slot. In this phase, the transmitter sends 2 information symbols for each of $\mathrm{Rx}_1, \mathrm{Rx}_2, \ldots, \mathrm{Rx}_{K-1}$, denoted by $(\mathbf{s}_1^1, \mathbf{s}_1^2)$, $(\mathbf{s}_2^1, \mathbf{s}_2^2)$, $\ldots$, $(\mathbf{s}_{K-1}^1, \mathbf{s}_{K-1}^2)$, along with one symbol, denoted by $\mathbf{s}_K$, for the K-th receiver. Let $\vec{\mathbf{g}}_{\mathcal{S}}(1)^{\perp}$, where $\mathcal{S} \subseteq \{1, \ldots, K-1\}$, denote a full row rank matrix of size $(K-|\mathcal{S}|) \times K$, where each row of $\vec{\mathbf{g}}_{\mathcal{S}}(1)^{\perp}$ is perpendicular to any $\vec{\mathbf{g}}_i(1)$ where $i \in \mathcal{S}$. We need to deliver one equation about $(\mathbf{s}_i^1, \mathbf{s}_i^2)$ interference-free to

Rx_i, for $i = 1, \ldots, K-1$. To this aim, the transmit signal at time 1 will be:

$$\vec{\mathbf{x}}_1(1) = \sum_{i=1}^{K-1} [\vec{\mathbf{g}}_{\{1,\ldots,K-1\}\setminus\{i\}}(1)^{\perp}]^{\top} \begin{bmatrix} \mathbf{s}_i^1 \\ \mathbf{s}_i^2 \end{bmatrix} + [\vec{\mathbf{g}}_{\{1,\ldots,K-1\}}(1)^{\perp}]^{\top} \mathbf{s}_K. \tag{C.43}$$

As a result, each of the first $K-1$ receivers obtain one equation in 2 desired symbols:

$$\mathbf{y}_i(1) = \vec{\mathbf{g}}_i(1)[\vec{\mathbf{g}}_{\{1,\ldots,K-1\}\setminus\{i\}}(1)^{\perp}]^{\top} \begin{bmatrix} \mathbf{s}_i^1 \\ \mathbf{s}_i^2 \end{bmatrix}, \qquad i = 1, \ldots, K-1; \tag{C.44}$$

and receiver K obtains $\mathbf{s}_K$ along with interference from the other symbols:

$$\mathbf{y}_K(1) = \sum_{i=1}^{K-1} \vec{\mathbf{g}}_K(1)[\vec{\mathbf{g}}_{\{1,\ldots,K-1\}\setminus\{i\}}(1)^{\perp}]^{\top} \begin{bmatrix} \mathbf{s}_i^1 \\ \mathbf{s}_i^2 \end{bmatrix} + \vec{\mathbf{g}}_K(1)[\vec{\mathbf{g}}_{\{1,\ldots,K-1\}}(1)^{\perp}]^{\top} \mathbf{s}_K, \tag{C.45}$$

which can be re-written as:

$$\mathbf{y}_K(1) = L_1(\mathbf{s}_1^1, \mathbf{s}_1^2) + L_2(\mathbf{s}_2^1, \mathbf{s}_2^2) + \ldots + L_{K-1}(\mathbf{s}_{K-1}^1, \mathbf{s}_{K-1}^2) + \left(\vec{\mathbf{g}}_K(1)[\vec{\mathbf{g}}_{\{1,\ldots,K-1\}}(1)^{\perp}]^{\top}\right) \mathbf{s}_K, \tag{C.46}$$

where $L_i(\mathbf{s}_i^1, \mathbf{s}_i^2) = \vec{\mathbf{g}}_K(1)[\vec{\mathbf{g}}_{\{1,\ldots,K-1\}\setminus\{i\}}(1)^{\perp}]^{\top} \begin{bmatrix} \mathbf{s}_i^1 \\ \mathbf{s}_i^2 \end{bmatrix}$. We observe that the K-th receiver has obtained 1 equation, and this equation has $(K-1)$ interfering order-2 symbols, where each order-2 symbol is desirable by one of the other $(K-1)$ receivers. In particular, each order-2 symbol $L_i(\mathbf{s}_i^1, \mathbf{s}_i^2)$ is desired by Rx_i.

The purpose of subsequent phases of the scheme is the following: in each phase i, we deliver the interference symbols of phase $i-1$ to the intended receivers while simultaneously sending new information symbols. This should be done in an iterative manner to create a new set of equations at the K-th receiver with net interference from a smaller set of receivers, where the interference is useful for that set of receivers. With this broad goal in mind, we next describe the transmission strategy for the general phase i.

C.7.2 Phase i

Duration of Phase i is $\binom{K-1}{i}$ time slots. Let us index the slots as $j = 1, 2, \ldots, \binom{K-1}{i}$.

Transmission in slot j**,** $j = 1, 2, \ldots, \binom{K-1}{i}$

In each time slot j, the transmitter selects i receivers out of first $(K-1)$ receivers. This splits the set of $(K-1)$ receivers into two disjoint sets, and for simplicity we denote these as:

- $\mathcal{R}$ (Repetition set): this is a set of i receivers. Let us denote the indices of the receivers in this set by $(p_1, p_2, \ldots, p_i)$.
- $\mathcal{F}$ (Fresh set): this is the remaining set of $(K-1-i)$ receivers, and we denote this set of receivers as $(p_{i+1}, \ldots, p_{K-1})$

The basic idea behind the scheme can now be explained clearly:

- Note that in phase $(i-1)$, the K-th receiver has obtained $\binom{K-1}{i-1}$ equations, where each equation is a linear combination of $(K-i)$ undesired symbols and the intended symbol (of the K-th receiver).
- Via delayed CSIT, the transmitter can reconstruct all of these equations within noise distortion.
- Out of these $\binom{K-1}{i-1}$ equations, the transmitter focuses on those equations which consist of all symbols from the receivers $p_{i+1}, \ldots, p_{K-1}$ (i.e., the receivers belonging to the fresh set $\mathcal{F}$). In total, there are exactly $\binom{i}{1} = i$ such equations. The reason is that each equation in phase $(i-1)$ has interference from exactly $(K-i)$ receivers. We zoom in on such equations with

interference from $(K-1-i)$ receivers $p_{i+1}, \ldots, p_{K-1}$, and thus the remaining flexibility is to choose 1 more interference symbol. The total remaining receivers to select from are $(K-1)-(K-1-i) = i$ and hence the number of ways is $\binom{i}{1} = i$.

- From each of these i equations, the transmitter reconstructs the only symbol in the equation which is desired by one of the receivers in the repetition set $(p_1, p_2, \ldots, p_i)$. Let us denote the reconstructed symbols by $\mathbf{s}_{p_1}(j), \ldots, \mathbf{s}_{p_i}(j)$. Also, we denote those i equations as following:

$$\mathbf{s}_{p_1}(j) + LC_1 : \text{ where } LC_1 \text{ is a linear combination of symbols for receivers in set } \mathcal{F} \cup \{K\} \tag{C.47}$$

$$\mathbf{s}_{p_2}(j) + LC_2 : \text{ where } LC_2 \text{ is a linear combination of symbols for receivers in set } \mathcal{F} \cup \{K\} \tag{C.48}$$

$$\vdots$$

$$\mathbf{s}_{p_i}(j) + LC_i : \text{ where } LC_i \text{ is a linear combination of symbols for receivers in set } \mathcal{F} \cup \{K\}. \tag{C.49}$$

- For each of the $(K-1-i)$ receivers in the fresh set, the transmitter sends 2 precoded fresh (i.e. new) symbols. Let us denote these as $(\mathbf{s}^1_{p_{i+1}}(j), \mathbf{s}^2_{p_{i+1}}(j))$, $(\mathbf{s}^1_{p_{i+2}}(j), \mathbf{s}^2_{p_{i+2}}(j))$ up to $(\mathbf{s}^1_{p_{K-1}}(j), \mathbf{s}^2_{p_{K-1}}(j))$.

Hence in the j-th slot of phase i, the transmitter sends:

$$\vec{\mathbf{x}}_i(j) = \underbrace{\sum_{r=1}^{i} [\vec{\mathbf{g}}_{\{1,\ldots,K-1\}\setminus\{p_r(j)\}}(j)^{\perp}]^{\top} \begin{bmatrix} \mathbf{s}_{p_r}(j) \\ 0 \end{bmatrix}}_{i \text{ repetition symbols}} + \underbrace{\sum_{r=i+1}^{K-1} [\vec{\mathbf{g}}_{\{1,\ldots,K-1\}\setminus\{p_r(j)\}}(j)^{\perp}]^{\top} \begin{bmatrix} \mathbf{s}^1_{p_r}(j) \\ \mathbf{s}^2_{p_r}(j) \end{bmatrix}}_{2(K-1-i) \text{ fresh symbols}}. \tag{C.50}$$

Clearly, the set of repetition receivers $\{p_1, p_2, \ldots, p_i\}$ receive one symbol without interference. Similarly, the set of receivers $\{p_{i+1}, \ldots, p_{K-1}\}$ also receive one clean (interference-free) useful symbol in this slot (which is a linear combination of the two fresh symbols).

Operation at Rx_K in time slot j of phase i

Let us now focus on Rx_K at the j-th time slot of phase i. Rx_K obtains

$$\mathbf{y}_K(i,j) = \sum_{r=1}^{i} \alpha_r(i,j)\mathbf{s}_{p_r}(j) + \sum_{r=i+1}^{K-1} LC(\mathbf{s}^1_{p_r}(j), \mathbf{s}^2_{p_r}(j)), \tag{C.51}$$

where $\alpha_r(i,j)$ denotes the coefficient of the symbol $\mathbf{s}_{p_r}(j)$ when received at Rx_K; and $LC(\mathbf{s}^1_{p_r}(j), \mathbf{s}^1_{p_r}(j))$ denotes the linear combination of $\mathbf{s}^1_{p_r}(j), \mathbf{s}^1_{p_r}(j)$ received at Rx_K. Note that from phase $(i-1)$, the receiver also has i equations $\mathbf{s}_{p_1}(j) + LC_1, \ldots, \mathbf{s}_{p_i}(j) + LC_i$ as mentioned in (C.47)-(C.49). Using these i equations together with (C.51), receiver K eliminates the i symbols $\mathbf{s}_{p_1}(j), \mathbf{s}_{p_2}(j), \ldots, \mathbf{s}_{p_i}(j)$; and it is left with an equation of the following form:

$$LC_{p_{i+1}}(j) + LC_{p_{i+2}}(j) + \ldots + LC_{p_{K-1}}(j) + \mathbf{s}_K. \tag{C.52}$$

This equation consists of $(K-1-i)$ interfering symbols, where each interfering symbol $LC_{p_r}(j)$ is desired by Rx_{p_r}, for $r = i+1, i+2, \ldots, (K-1)$.

Recall that the slot index j varies from 1 to $\binom{K-1}{i}$, each slot corresponding to the partitioning of the set of $K-1$ receivers into two disjoint sets of size i and $(K-1-i)$. Each slot gives the K-th receiver one equation with interference from exactly $(K-1-i)$ receivers. Hence, in total, at the end of phase i, the receiver has $\binom{K-1}{i}$ equations, and each equation has interference from symbols desired by exactly $(K-1-i)$ receivers. Thus, we can now readily apply this process iteratively.

Phase $K-1$ (the last phase; corresponding to $i = K-1$)

Before the last phase $K-1$, (i.e., just after phase $K-2$), the K-th receiver has $\binom{K-1}{i-1} = \binom{K-1}{K-2} = K-1$ equations, and each equation has interference from exactly $(K-1)-(i-1) = (K-1)-(K-2) = 1$ receiver. Hence, the K-th receiver has $K-1$ equations of the following form before the last phase:

$$LC'_1 + \mathbf{s}_K, LC'_2 + \mathbf{s}_K, \ldots, LC'_{K-1} + \mathbf{s}_K, \tag{C.53}$$

where LC'_1 is desired by receiver 1, LC'_2 is desired by receiver 2, etc.

In the last phase, whose duration is only 1 slot (since $\binom{K-1}{K-1} = 1$), the transmitter sends $LC'_1, \ldots, LC'_{K-1}$ without any interference to receivers $1, \ldots, K-1$ by utilizing instantaneous CSIT. Receiver K obtains a linear combination of $LC'_1, \ldots LC'_{K-1}$. Hence, the Kth receiver has K equations in K variables $LC'_1, LC'_2, \ldots, LC'_K$ and s_K. Therefore, it can decode s_K; and the proof is complete.

C.7.3 Illustrative Example – 4 User MISO BC

Here, we present the achievable scheme for $K = 4$ to clearly illustrate the idea behind the iterative scheme. For the case of 4-user MISO BC with $PPPD$, the goal is to achieve:

$$(d_1, d_2, d_3, d_4) = \left(1, 1, 1, \frac{1}{2^3}\right). \tag{C.54}$$

Here, the scheme has $K = 4$ phases, with the following phase durations:

- Phase 0: $\binom{3}{0} = 1$ time slots; Tx sends two new information symbols for each of the first three receivers, and one symbol for the fourth receiver. Each of

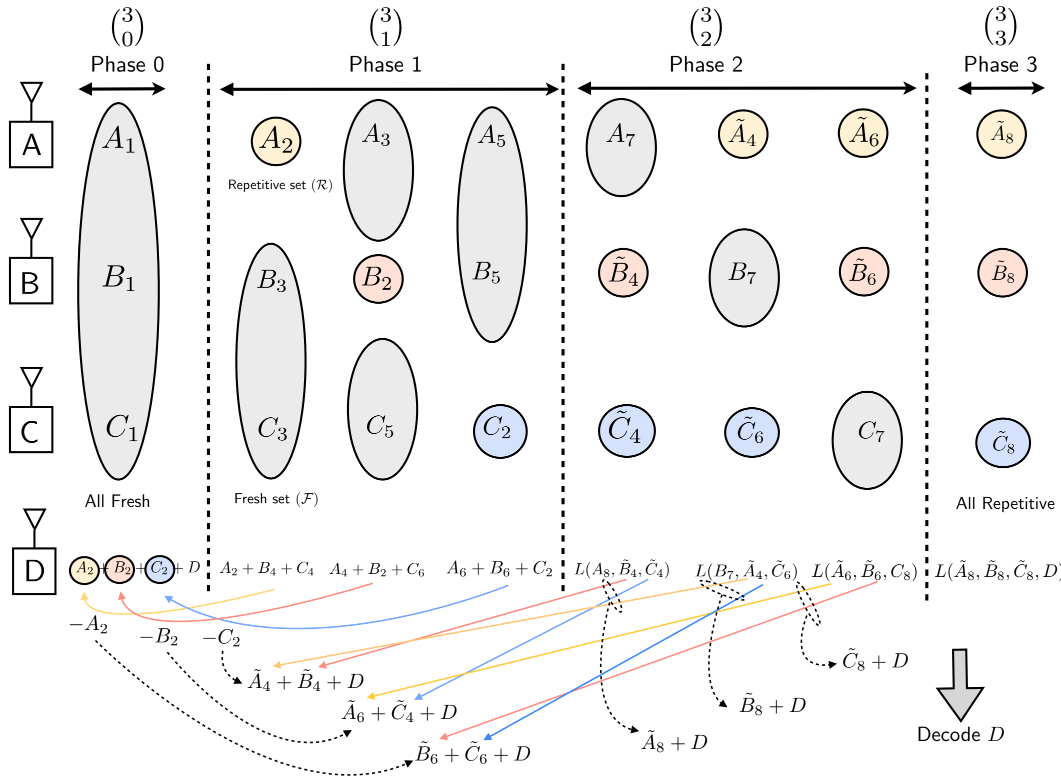


Figure C.1: Scheme for 4-user MISO BC: *PPPD* Setting.

the first three receivers will receive a linear combination of its two desired symbols without any interference.

- Phase 1: $\binom{3}{1}$ = 2 time slots; in each time slot, Tx sends the signal received in the past by Rx_K with respect to the symbols of one of the first three receivers; and it also sends two new information symbols for each of the other 2 receivers supplying instantaneous CSIT.
- Phase 2: $\binom{3}{2}$ = 3 time slots; in each slot, Tx sends fresh information for 1 receiver with instantaneous CSIT and supplies past signals received by Rx_K with respect to the remaining 2 receivers supplying instantaneous CSIT.
- Phase 3: $\binom{3}{3}$ = 1 time slot; Tx sends past received signals by Rx_K which are desired by the three receivers supplying instantaneous CSIT.

See Figure C.1 which illustrates the achievable scheme for 4-user MISO BC, where the first 3 receivers supply instantaneous CSIT, while the fourth receiver supplies delayed CSIT.

C.8 Proof of Claim 2

We first re-state Claim 2 here for convenience.

Claim 2.

$$\sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{m_j(n)}{2^j} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]. \tag{C.55}$$

To prove Claim 2 we first prove the following inequality by induction, and then show how it leads to proving Claim 2.

$$\sum_{j=1}^{i-1} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_i \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2^{i-1}} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]],$$

$$i = 2, \dots, |\mathcal{P}| + |\mathcal{D}| - 1. \tag{C.56}$$

We prove (C.56) by induction on i. For the base case of $i = 2$, the inequality in (C.56) simplifies to

$$\frac{m_1(n)}{2} + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_2 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2} \overset{a.s.}{\leq} \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]. \tag{C.57}$$

Hence, the base case of $i = 2$ holds due to Lemma 17 and (4.9). Suppose that the induction hypothesis is true for $i = s$. We show that it will also hold for $i = s + 1$. By our assumption we have

$$\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]$$

$$
\begin{aligned}
&\overset{a.s.}{\geq} \sum_{j=1}^{s-1} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_s \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2^{s-1}} \\
&\overset{\substack{(\text{Lemma 17})\\ a.s.}}{\geq} \sum_{j=1}^{s-1} \frac{m_j(n)}{2^j} + \frac{\frac{\text{rank}[\mathbf{G}^n_s[\mathbf{V}^n_s \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]] - \text{rank}[\mathbf{G}^n_s[\mathbf{V}^n_{s+1} \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]] + \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_{s+1} \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2}}{2^{s-1}} \\
&\overset{\substack{(\text{Lemma 2})\\ a.s.}}{\geq} \sum_{j=1}^{s-1} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_s[\mathbf{V}^n_1 \dots \mathbf{V}^n_k]] - \text{rank}[\mathbf{G}^n_s[\cup_{\substack{i\in\{1,\dots,k\}\\ i\neq s}} \mathbf{V}^n_i]]}{2^s} \\
&\quad + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_{s+1} \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2^s} \\
&\overset{\substack{(4.9)\\ a.s.}}{=} \sum_{j=1}^{s-1} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_s \mathbf{V}^n_s] + \text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_{s+1} \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2^s} \\
&\overset{\substack{(4.9)\\ a.s.}}{=} \sum_{j=1}^{s} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_{s+1} \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]]}{2^s}.
\end{aligned}
$$

Hence, the induction hypothesis holds for $i = s + 1$ as well; and as a result, the proof of (C.56) is complete. We now show how (C.56) leads to proof of Claim 2. Let $i = |\mathcal{P}| + |\mathcal{D}| - 1$. Then, by (C.56),

$$
\begin{aligned}
\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}[\mathbf{V}^n_1 \dots \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]] &\overset{a.s.}{\geq} \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-2} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|} \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]}{2^{|\mathcal{P}|+|\mathcal{D}|-2}} \\
&\overset{\substack{(\text{Lemmaa 18})\\ a.s.}}{\geq} \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-2} \frac{m_j(n)}{2^j} \\
&\quad + \frac{\text{rank}[[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|-1}; \mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|}]\mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]}{2^{|\mathcal{P}|+|\mathcal{D}|-1}} \\
&\geq \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-2} \frac{m_j(n)}{2^j} + \frac{\text{rank}[\mathbf{G}^n_{|\mathcal{P}|+|\mathcal{D}|-1} \mathbf{V}^n_{|\mathcal{P}|+|\mathcal{D}|-1}]}{2^{|\mathcal{P}|+|\mathcal{D}|-1}} \\
&\overset{\substack{(4.9)\\ a.s.}}{=} \sum_{j=1}^{|\mathcal{P}|+|\mathcal{D}|-1} \frac{m_j(n)}{2^j},
\end{aligned}
$$

which completes the proof of Claim 2.

C.9 Proof of Claim 3

We first re-state the Claim for convenience.

Claim 3.

$$\frac{\text{rank}[[\mathbf{G}_1^n;\ldots;\mathbf{G}_k^n][\mathbf{V}_1^n\ldots\mathbf{V}_{|\mathcal{P}|}^n]]}{k} \overset{a.s.}{\leq} \frac{\text{rank}[[\mathbf{G}_{|\mathcal{P}|+1}^n;\ldots;\mathbf{G}_{|\mathcal{P}|+|\mathcal{D}|}^n][\mathbf{V}_1^n\ldots\mathbf{V}_{|\mathcal{P}|}^n]]}{|\mathcal{D}|}. \tag{C.58}$$

Proof. We consider the notations (C.11), (C.12); and we use $\mathbf{Y}_j^n \triangleq \mathbf{G}_j^n[\mathbf{V}_1^n\ldots\mathbf{V}_{|\mathcal{P}|}^n]$, and $\mathbf{Y}_j(t) \triangleq \vec{\mathbf{g}}_j(t)[\mathbf{V}_1(t)\ldots\mathbf{V}_{|\mathcal{P}|}(t)]$. Furthermore, we denote by $\mathbf{Y}_{\mathcal{S}}^n$ the column concatenation of matrices $\mathbf{G}_j^n[\mathbf{V}_1^n\ldots\mathbf{V}_{|\mathcal{P}|}^n]$, where $j \in \mathcal{S}$. Therefore, we need to show that

$$\frac{\text{rank}[\mathbf{Y}_1^n;\ldots;\mathbf{Y}_k^n]}{k} \overset{a.s.}{\leq} \frac{\text{rank}[\mathbf{Y}_{\mathcal{D}}^n]}{|\mathcal{D}|}. \tag{C.59}$$

For all $t = 1,\ldots,n$, we have

$$\begin{aligned}
&(|\mathcal{P}|+|\mathcal{N}|)\times\text{rank}[\mathbf{Y}_{\mathcal{D}}(t)|\mathbf{Y}_{\mathcal{D}}^{t-1}]\\
&\overset{(C.11)}{=}\sum_{i=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|}(|\mathcal{P}|+|\mathcal{N}|)\times\text{rank}[\mathbf{Y}_i(t)|[\mathbf{Y}_{\mathcal{D}}^{t-1};\mathbf{Y}_{|\mathcal{P}|+1}(t);\ldots;\mathbf{Y}_{i-1}(t)]]\\
&\overset{(a)}{\underset{a.s.}{\geq}}\sum_{i=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|}\sum_{j\in\mathcal{P}\cup\mathcal{N}}\text{rank}[\mathbf{Y}_j(t)|[\mathbf{Y}_{\mathcal{D}}^{t-1};\mathbf{Y}_{|\mathcal{P}|+1}(t);\ldots;\mathbf{Y}_{i-1}(t)]]\\
&\overset{(C.12)}{\geq}\sum_{i=|\mathcal{P}|+1}^{|\mathcal{P}|+|\mathcal{D}|}\sum_{j\in\mathcal{P}\cup\mathcal{N}}\text{rank}[\mathbf{Y}_j(t)|[\mathbf{Y}_{\mathcal{D}}^n;\mathbf{Y}_{\mathcal{P}\cup\mathcal{N}}^{t-1}]] \quad = \quad |\mathcal{D}|\sum_{j\in\mathcal{P}\cup\mathcal{N}}\text{rank}[\mathbf{Y}_j(t)|[\mathbf{Y}_{\mathcal{D}}^n;\mathbf{Y}_{\mathcal{P}\cup\mathcal{N}}^{t-1}]]\\
&\overset{(C.13)}{\geq}|\mathcal{D}|\times\text{rank}[\mathbf{Y}_{\mathcal{P}\cup\mathcal{N}}(t)|[\mathbf{Y}_{\mathcal{D}}^n;\mathbf{Y}_{\mathcal{P}\cup\mathcal{N}}^{t-1}]],
\end{aligned} \tag{C.60}$$

where (a) follows from the same arguments as in (C.14)-(C.18) which were used to show that

$$\text{rank}[\mathbf{Y}_i(t)|[\mathbf{Y}^{t-1};\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{i-1}(t)]] \overset{a.s.}{\geq} \text{rank}[\mathbf{Y}_{j+1}(t)|[\mathbf{Y}^{t-1};\mathbf{Y}_1(t);\ldots;\mathbf{Y}_{i-1}(t)]],$$

for the case where $i \in \{1,\ldots,j\} \subseteq \mathcal{D}$, and $\mathbf{Y}^{t-1} \triangleq [\mathbf{Y}_1^{t-1};\ldots;\mathbf{Y}_j^{t-1}]$.

By summing both sides of the inequality (C.60) over all $t = 1,\ldots,n$, we obtain

$$(|\mathcal{P}|+|\mathcal{N}|)\times\text{rank}[\mathbf{Y}_{\mathcal{D}}^n] \overset{a.s.}{\geq} |\mathcal{D}|\times\text{rank}[\mathbf{Y}_{\mathcal{P}\cup\mathcal{N}}^n|\mathbf{Y}_{\mathcal{D}}^n]$$

$$\stackrel{(C.11)}{=} |\mathcal{D}| \times \text{rank}[\mathbf{Y}_1^n; \ldots; \mathbf{Y}_k^n] - |\mathcal{D}| \times \text{rank}[\mathbf{Y}_{\mathcal{D}}^n]. \qquad \text{(C.61)}$$

Finally, by rearranging the above inequality we obtain (C.59), which proves Claim 3. $\square$

APPENDIX D

APPENDIX FOR CHAPTER 5

D.1 Proof of Tightness of the Bounds in Theorem 1

We prove that the upper and lower bounds given in (5.8) are tight. More specifically, we show that there exist N, M, and some channel success probabilities for which C_{T^3} gets arbitrarily close to $C_{\text{det}} + N$. In addition, there exist N, M, and some channel success probabilities for which $O(|C_{\mathsf{T}^3} - C_{\text{det}}|) = O(\sqrt{NC_{\text{det}}})$.

D.1.1 Proof of Tightness of the Upper Bound

We show that for any given N and $0 < \epsilon < N$ there exist M, τ, and channel success probabilities such that $C_{\mathsf{T}^3} - C_{\text{det}} = N - \epsilon$. We set $M = N\tau$, and we choose C_{det} such that $C_{\text{det}} < M - N$ and $\frac{C_{\text{det}}}{N} \in \mathbb{N}$. Further, for the channel between AP_i and Rx_j we set the channel success probability $p_{ij} = \frac{C_{\text{det}}+N-\epsilon}{N\tau}$, where $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$. Therefore, according to symmetry, both the optimal assignment which results in C_{T^3} and the optimal assignment for the relaxed problem which results in C_{det} assign τ packets to each AP. Furthermore, without loss of generality we can assume that for AP_i packets of clients $j = 1 + (i-1)\tau, \ldots, i\tau$ are assigned to AP_i. It is easy to check that the following inequalities hold for any AP_i, $i = 1, 2, \ldots, N$:

$$\sum_{j=1+(i-1)\tau}^{1+(i-1)\tau+C_{\text{det}}/N} \frac{1}{p_{ij}} = (\frac{C_{\text{det}}}{N})(\frac{N\tau}{C_{\text{det}} + N - \epsilon}) < \tau,$$

$$\tau < \sum_{j=1+(i-1)\tau}^{1+(i-1)\tau+C_{\text{det}}/N+1} \frac{1}{p_{ij}}.$$

Therefore, the maximum number of packets that can be packed in the relaxed problem is C_{det}. Now, we calculate the expected number of packet deliveries: For any AP_i the expected number of successful deliveries during one interval is $\tau(\frac{C_{\text{det}}+N-\epsilon}{N\tau}) = \frac{C_{\text{det}}+N-\epsilon}{N}$. Therefore, we have $C_{\mathsf{T}^3} = N(\frac{C_{\text{det}}+N-\epsilon}{N}) = C_{\text{det}} + N - \epsilon$. Hence, $C_{\mathsf{T}^3} - C_{\text{det}} = N - \epsilon$.

D.1.2 Proof of Tightness of the Order of the Lower Bound

We show that there exists a wireless network realization for which $O(|C_{\mathsf{T}^3} - C_{\text{det}}|) = O(\sqrt{NC_{\text{det}}})$. More specifically, for a given N we show that there exist a positive constant k along with M, τ, such that $C_{\text{det}} - C_{\mathsf{T}^3} > k\sqrt{NC_{\text{det}}}$. We choose C_{det} such that $\frac{C_{\text{det}}}{N} \in \mathbb{N}$, and we set $M = C_{\text{det}}$. In addition, we set the channel success probability $p_{ij} = p = \frac{C_{\text{det}}}{N\tau} < 1$ for some $\tau \in \mathbb{N}$, where $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$.

Therefore, both the optimal assignment which results in C_{T^3} and the optimal assignment for the relaxed problem which results in C_{det} assign $\frac{C_{\text{det}}}{N}$ packets to each AP. It is easy to check that our chosen C_{det} is actually the solution to the relaxed problem. Now, let Y denote the number of successful deliveries for one of the AP's. Thus, $C_{\mathsf{T}^3} = NE[Y]$. Also, let l denote the number of packets that can be packed in a bin corresponding to a certain AP. Therefore, $l = \frac{C_{\text{det}}}{N}$, and $p = \frac{l}{\tau}$. We only need to show that there exists a constant k such that $l - E[Y] > k\sqrt{l}$. Noting that $l = p\tau$ we have

$$
\begin{aligned}
l - E[Y] =& p\tau - [\sum_{j=1}^{l} j\binom{\tau}{j}p^j(1-p)^{\tau-j} + l\sum_{j=l+1}^{\tau}\binom{\tau}{j}p^j(1-p)^{\tau-j}] \\
=& \sum_{j=1}^{\tau} j\binom{\tau}{j}p^j(1-p)^{\tau-j} - [\sum_{j=1}^{l} j\binom{\tau}{j}p^j(1-p)^{\tau-j} + l\sum_{j=l+1}^{\tau}\binom{\tau}{j}p^j(1-p)^{\tau-j}]
\end{aligned}
$$

$$
\begin{aligned}
=& p\tau \sum_{j=l+1}^{\tau} \binom{\tau-1}{j-1} p^{j-1}(1-p)^{\tau-j} - l \sum_{j=l+1}^{\tau} \binom{\tau}{j} p^{j}(1-p)^{\tau-j} \\
=& l[\sum_{j=l+1}^{\tau} \binom{\tau-1}{j-1} p^{j-1}(1-p)^{\tau-j} - \sum_{j=l+1}^{\tau} (\binom{\tau-1}{j} + \binom{\tau-1}{j-1}) p^{j}(1-p)^{\tau-j}] \\
=& l[\sum_{j=l+1}^{\tau} \binom{\tau-1}{j-1} p^{j-1}(1-p)^{\tau-j+1} - \sum_{j=l+1}^{\tau-1} \binom{\tau-1}{j} p^{j}(1-p)^{\tau-j}] \\
=& l\binom{\tau-1}{l} p^{l}(1-p)^{\tau-l}.
\end{aligned}
$$

Now note that $\binom{\tau-1}{l} = \frac{(\tau-1)!}{l!(\tau-1-l)!} = \frac{\tau-l}{\tau}\binom{\tau}{l} = (1-p)\binom{\tau}{l}$. Therefore, $l\binom{\tau-1}{l}p^l(1-p)^{\tau-l} = \tau\binom{\tau}{l}p^{l+1}(1-p)^{\tau-l+1}$. By Theorem 2.6 of [79] we know that for positive integers m,n,q, with $m > q \geq 1$ and $n \geq 1$

$$
\binom{mn}{qn} > \frac{1}{\sqrt{2\pi}} e^{\frac{1}{12n}(\frac{1}{m}-\frac{1}{q}-\frac{1}{m-q})} n^{-\frac{1}{2}} \frac{m^{mn+\frac{1}{2}}}{(m-q)^{(m-q)n+\frac{1}{2}} q^{qn+\frac{1}{2}}}.
$$

Substituting n by 1, m by τ, and q by l we get:

$$
\begin{aligned}
\binom{\tau}{l} &> \frac{1}{\sqrt{2\pi}} \frac{\tau^{\tau+\frac{1}{2}}}{(\tau-l)^{(\tau-l)+\frac{1}{2}} l^{l+\frac{1}{2}}} e^{\frac{1}{12}(\frac{1}{\tau}-\frac{1}{l}-\frac{1}{\tau-l})} \\
&= \frac{1}{\sqrt{2\pi}} \frac{\tau^{\tau+\frac{1}{2}}}{(\tau(1-p))^{\tau-l+\frac{1}{2}} (p\tau)^{l+\frac{1}{2}}} e^{\frac{1}{12}(\frac{1}{\tau}-\frac{1}{l}-\frac{1}{\tau-l})} \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\tau p(1-p)}} \frac{1}{p^l(1-p)^{\tau-l}} e^{\frac{1}{12}(\frac{1}{\tau}-\frac{1}{l}-\frac{1}{\tau-l})}.
\end{aligned}
$$

However, $\frac{1}{\tau} - \frac{1}{l} - \frac{1}{\tau-l} > -2$. Therefore, $\binom{\tau}{l} > \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\tau p(1-p)}} \frac{1}{p^l(1-p)^{\tau-l}} e^{-\frac{1}{6}}$. Hence, we get

$$
\begin{aligned}
l - E[Y] &= l\binom{\tau-1}{l} p^l(1-p)^{\tau-l} \\
&= \tau\binom{\tau}{l} p^{l+1}(1-p)^{\tau-l+1} \\
&> \tau \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\tau p(1-p)}} \frac{1}{p^l(1-p)^{\tau-l}} e^{-\frac{1}{6}} p^{l+1}(1-p)^{\tau-l+1} \\
&> e^{-\frac{1}{6}} \sqrt{\frac{1-p}{2\pi}} \sqrt{l}.
\end{aligned}
$$

Thus, by setting $k = e^{-\frac{1}{6}} \sqrt{\frac{1-p}{2\pi}}$ the proof will be complete.

D.2 Proof of Lemma 20

Lemma 20 states that C_{T^3} can be achieved using a greedy static scheduling policy.

Proof. The proof consists of two parts. In part A we prove that when looking at class of scheduling policies that use the same assignment of packets to AP's for all intervals, the maximal T^3, R^*, can be achieved using a greedy static scheduling policy. In part B we prove that no policy in general can achieve any T^3 greater than R^*. Considering these two parts together, the desired result will be obtained.

D.2.1 Proving that maximal T^3 over the class of scheduling policies that use the same assignment of packets for all intervals, is achieved using a greedy static scheduling policy:

There are a total of N^M different possible ways of assigning packets to AP's for each interval. We enumerate these different assignment policies by $i = 1, 2, \ldots, N^M$. For an assignment i, $i \in \{1, 2, \ldots, N^M\}$, we define $R(i)$ to be the supremum of achievable total timely throughputs, given that the assignment i is used for all intervals.

Define $R^* \triangleq \max_{i\in\{1,2,\ldots,N^M\}} R(i)$. We will now prove that there is a greedy static policy which achieves R^*. It is sufficient to show that for all $i \in \{1, 2, \ldots, N^M\}$

$R(i)$ can be achieved using a greedy static policy. Consider an arbitrary i, $i \in \{1, 2, \ldots, N^M\}$. Since the set of packets assigned to different AP's are disjoint, and their channels are independent of each other, $R(i)$ is just the summation of supremums of timely throughputs on different AP's, when assignment i is used for all intervals.

The result in [33] states that the timely throughput region for each AP is a scaled version of a polymatroid (i.e., a polymatroid that has each of its coordinates scaled by a constant factor). Moreover, in [96] it has been shown that each of the corner points of this polytope can be achieved using a static policy. Therefore, when assignment i is used, there is a static policy which achieves $R(i)$.

Furthermore, when using a static policy, according to LLN the resulting T^3 is equal to expected number of deliveries during one interval for that static policy. So, $R(i)$ is the highest expected number of deliveries among static scheduling policies that use assignment policy i.

The following lemma implies that the highest expected number of deliveries among the static policies that use the same assignment policy is achieved by the one which serves the packets based on their channel success probabilities, and in decreasing order. To prove this, it is sufficient to prove that for any given order if we swap the order of two adjacent clients in such a way that the client with the higher corresponding p_{ij} is prioritized higher, then the expected number of deliveries will be no less than before swapping.

Lemma 32. *Let $\tau \in \mathbb{N}$ and $G_1, G_2, \ldots, G_q$ be independent geometric random variables with parameters $p_1, p_2, \ldots, p_q$, respectively. Suppose that $p_d < p_{d+1}$ for some $d \in \{1, 2, \ldots, q-1\}$. In addition, let $G'_1, G'_2, \ldots, G'_q$ be independent geometric random variables (and independent of G_i's) with parameters $p_1, p_2, \ldots, p_{d-1}, p_{d+1}, p_d, p_{d+2}, \ldots, p_q$,* *respectively. Then,*

$$\sum_{i=1}^{q} \Pr(\sum_{j=1}^{i} G_j \leq \tau) \leq \sum_{i=1}^{q} \Pr(\sum_{j=1}^{i} G'_j \leq \tau).$$

Proof. We have

$$\begin{aligned}
\sum_{i=1}^{q} \Pr(\sum_{j=1}^{i} G_j \leq \tau) &= \sum_{i=1}^{d-1} \Pr(\sum_{j=1}^{i} G_j \leq \tau) + \Pr(\sum_{j=1}^{d} G_j \leq \tau) + \sum_{i=d+1}^{q} \Pr(\sum_{j=1}^{i} G_j \leq \tau) \\
&= \sum_{i=1}^{d-1} \Pr(\sum_{j=1}^{i} G'_j \leq \tau) + \Pr(G_d + \sum_{j=1}^{d-1} G'_j \leq \tau) + \sum_{i=d+1}^{q} \Pr(\sum_{j=1}^{i} G'_j \leq \tau) \\
&\overset{(a)}{\leq} \sum_{i=1}^{d-1} \Pr(\sum_{j=1}^{i} G'_j \leq \tau) + \Pr(G'_d + \sum_{j=1}^{d-1} G'_j \leq \tau) + \sum_{i=d+1}^{q} \Pr(\sum_{j=1}^{i} G'_j \leq \tau) \\
&= \sum_{i=1}^{q} \Pr(\sum_{j=1}^{i} G'_j \leq \tau),
\end{aligned}$$

where (a) follows from the fact that success probability of G_d, which is p_d, is less than success probability of G'_d, which is p_{d+1}. □

Lemma 32 implies that when serving packets of some clients on an AP, one should serve them according to their channel success probabilities, and in decreasing order in order to maximize the expected number of deliveries. This is an intuitive fact, and Lemma 32 formalizes this fact. In conclusion, R^* can be achieved by a greedy static policy.

D.2.2 Proving that no policy in general can achieve any T^3 better than R^*:

Consider an arbitrary scheduling policy $\eta \in \mathcal{S}$ (not necessarily a static policy); we will show that $\mathsf{T}^3(\eta) \leq R^*$. Define the variable $N^i_j(k, \eta)$ to denote the outcome for client j using assignment i on interval k, i.e., if packet of client j is delivered during interval k when scheduling policy η and assignment i are used, $N^i_j(k,\eta) = 1$; otherwise, $N^i_j(k,\eta) = 0$. Moreover, define function U as a mapping which is used by η from intervals to assignment policies:

$$U : [\mathbb{N}, \mathcal{S}] \to \{1, 2, \dots, N^M\}.$$

Therefore, $U(k,\eta)$ is the assignment policy used by η for interval $k, k \in \mathbb{N}$. We call $\omega = \{U(k,\eta), N^i_j(k,\eta)\}_{k=1}^{\infty}$ an outcome for policy η over infinite intervals. In addition, we denote the set of all possible outcomes for policy η over infinite intervals by $\Omega(\eta)$.

In addition, define I to be the set of assignments that occur infinite times. More precisely,

$$I \triangleq \{i \in \{1, 2, \dots, N^M\} | \forall L \in \mathbb{N}, \exists T \in \mathbb{N} \quad s.t. \quad L \le \sum_{k=1}^{T} 1(U(k,\eta) = i)\}.$$

According to the definition of $\mathsf{T}^3(\eta)$ [1] there exists a subset of $\Omega(\eta)$, denoted by A, such that $P(A) = 1$ and for all $\omega = \{U(k,\eta), N^i_j(k,\eta)\}_{k=1}^{\infty}$ and $\omega \in A$,

$$\mathsf{T}^3(\eta) \le \limsup_{T\to\infty}\left(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M}\sum_{i=1}^{N^M} N^i_j(k,\eta)}{T}\right).$$

Therefore, for any outcome $\omega = \{U(k,\eta), N^i_j(k,\eta)\}_{k=1}^{\infty} \in A$, we have

$$\begin{aligned}
\mathsf{T}^3(\eta) &\le \limsup_{T\to\infty}\left(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M}\sum_{i=1}^{N^M} N^i_j(k,\eta)}{T}\right) \\
&\stackrel{(a)}{=} \limsup_{T\to\infty}\left(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M}\sum_{i\in I} N^i_j(k,\eta)}{T}\right) \\
&\stackrel{(b)}{=} \limsup_{T\to\infty}\left(\sum_{i\in I}\left(\frac{\sum_{k=1}^{T} 1(U(k,\eta) = i)}{T}\right) \times \left(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M} N^i_j(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta) = i)}\right)\right),
\end{aligned} \tag{D.1}$$

[1] $\mathsf{T}^3(\eta) = \sup \quad R \quad s.t. \quad \limsup_{T\to\infty} \frac{\sum_{k=1}^{T}\sum_{j=1}^{M}\sum_{i=1}^{N^M} N^i_j(k,\eta)}{T} \ge R$ with probability 1.

where (a) follows from the fact that the assignment i, where $i \notin I$, does not contribute to the value of limsup according to the definition of I. In addition, (b) is true because the fraction $\frac{\sum_{k=1}^{T}\sum_{j=1}^{M}\sum_{i\in I} N_j^i(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta)=i)}$ is properly defined for $i \in I$ since its denominator is not zero as $T \to \infty$. The reason why the denominator is not zero as $T \to \infty$ is that there exists $r \in \mathbb{N}$ such that $\sum_{k=1}^{r} 1(U(k,\eta) = i) \geq 1$ for $i \in I$ according to the definition of I. This means that for $T > r$, the fraction is well-defined.

Moreover, since $\limsup_{T\to\infty}(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M}\sum_{i\in I} N_j^i(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta)=i)})$ is the average number of successful deliveries for intervals for which assignment i is applied, there exists a subset of $\Omega(\eta)$, denoted by B, such that $P(B) > 0$ and for all $\omega = \{U(k,\eta), N_j^i(k,\eta)\}_{k=1}^{\infty}$, $\omega \in B$,

$$\limsup_{T\to\infty}(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M} N_j^i(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta) = i)}) \leq R^*.$$ [2]

In addition, note that $P(A \cap B) = P(A) - P(A \cup B) + P(B) = P(B) > 0$, which means $A \cap B$ is not empty. Hence, using (D.1) there is an outcome of η, $\omega = \{U(k,\eta), N_j^i(k,\eta)\}_{k=1}^{\infty}$ and $\omega \in A \cap B$, for which

$$\begin{aligned}
\mathsf{T}^3(\eta) &\leq \limsup_{T\to\infty}(\sum_{i\in I}(\frac{\sum_{k=1}^{T} 1(U(k,\eta) = i)}{T}) \times (\frac{\sum_{k=1}^{T}\sum_{j=1}^{M} N_j^i(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta) = i)})) \\
&\overset{(c)}{\leq} \limsup_{T\to\infty}(\sum_{i\in I} \frac{\sum_{k=1}^{T} 1(U(k,\eta) = i)}{T}) \times R^* \\
&\overset{(d)}{\leq} R^*,
\end{aligned}$$

where (c) follows from the fact that for $\omega = \{U(k,\eta), N_j^i(k,\eta)\}_{k=1}^{\infty}$, and $\omega \in A \cap B$, $\limsup_{T\to\infty}(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M} N_j^i(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta)=i)}) \leq R^*$, and also using Lemma 33. Finally (d) follows from

[2]This is true because if $\limsup_{T\to\infty}(\frac{\sum_{k=1}^{T}\sum_{j=1}^{M} N_j^i(k,\eta)}{\sum_{k=1}^{T} 1(U(k,\eta)=i)}) > R^*$ with probability 1, then we have a scheduling policy which uses the same assignment policy for all intervals, and achieves a T^3 which is strictly greater than R^* which is not possible according to the result in part A of the proof.

the fact that for each interval the scheduling policy can choose at most one of the N^M different possible assignments, or in other words, $\sum_{i\in I} \frac{\sum_{k=1}^{T} 1(U(k,\eta)=i)}{T} \leq 1$ for all $T \in \mathbb{N}$.

Therefore, for scheduling policy η, $\mathsf{T}^3(\eta) \leq R^*$. Using Part A and Part B we conclude that $C_{\mathsf{T}^3} = R^*$, and C_{T^3} can be achieved using a greedy static policy. $\square$

Below we provide Lemma 33 and its proof.

Lemma 33. *Suppose L is an integer, and $\{A_{1T}\}_{T=1}^{\infty}, \{A_{2T}\}_{T=1}^{\infty}, \ldots, \{A_{LT}\}_{T=1}^{\infty}$, and $\{B_{1T}\}_{T=1}^{\infty}, \{B_{2T}\}_{T=1}^{\infty}, \ldots, \{B_{LT}\}_{T=1}^{\infty}$ are non-negative real sequences, where* $\limsup_{T\to\infty} \sum_{i=1}^{L} A_{iT} < \infty$, *and for any $i \in \{1, 2, \ldots, L\}$,* $\limsup_{T\to\infty} B_{iT} \leq B$. *Then,*

$$\limsup_{T\to\infty} \sum_{i=1}^{L} A_{iT} B_{iT} \leq (\limsup_{T\to\infty} \sum_{i=1}^{L} A_{iT}) \times B.$$

Proof. Consider an arbitrary $\epsilon > 0$. Since $\forall i \in \{1, 2, \ldots, L\} \quad \limsup_{T\to\infty} B_{iT} \leq B$,

$$\exists M \in \mathbb{N}, \quad s.t. \quad \forall i \in \{1, 2, \ldots, L\}, T \geq M \quad B_{iT} \leq B + \epsilon.$$

Therefore, for all $r \geq M$ we will have $\sup_{T\geq r} \sum_{i=1}^{L} A_{iT} B_{iT} \leq \sup_{T\geq r} \sum_{i=1}^{L} A_{iT}(B+\epsilon)$. Hence, $\lim_{r\to\infty} \sup_{T\geq r} \sum_{i=1}^{L} A_{iT} B_{iT} \leq (B+\epsilon) \lim_{r\to\infty} \sup_{T\geq r} \sum_{i=1}^{L} A_{iT}$. Since the inequality is true for any $\epsilon > 0$, we have $\limsup_{T\to\infty} \sum_{i=1}^{L} A_{iT} B_{iT} \leq (\limsup_{T\to\infty} \sum_{i=1}^{L} A_{iT}) \times B$. $\square$

D.3 Proof of Lemma 21

For $l = 0$, we have $p_i < \frac{1}{\tau}$, for $i = 1, 2, \ldots, q$. Therefore, $E[Y]$ in this case is less than that of the case in which $p_1 = p_2 = \ldots = p_q = \frac{1}{\tau}$. On the other hand, for $p_1 = p_2 = \ldots = p_q = \frac{1}{\tau}$ $E[Y] \leq \tau \times \frac{1}{\tau} = 1$. Hence, the statement is true for $l = 0$. Now, suppose that $l > 0$. We know that $l = \max \quad \hat{l} \quad s.t. \quad \sum_{i=1}^{\hat{l}} 1/p_i \leq \tau$. Therefore, we have

$$\sum_{i=1}^{l} \frac{1}{p_i} \leq \tau < \sum_{i=1}^{l+1} \frac{1}{p_i}.$$

We will show that $E[Y]$ can be at most $l + 1$. Without loss of generality we can omit p_i's that are equal to zero; because by omitting them neither of $E[Y]$ nor l change, and $E[Y] - l$ would remain the same. So, we suppose that $1 \geq p_1 \geq p_2 \geq \ldots \geq p_q > 0$. It is sufficient to prove the lemma for the case of $q = \tau$; because if we have less than τ geometric random variables, $E[Y]$ will be less. On the other hand, we do not need to consider the case $q > \tau$; since for $i > \tau$, $\Pr(\sum_{j=1}^{i} G_j \leq \tau) = 0$. Therefore, we suppose that $q = \tau$.

Let $X_l = \sum_{i=1}^{l} G_i$, where $G_i = Geom(p_i)$. By this notation we have:

$$\Pr(X_l > \tau) = \sum_{i=0}^{l-1} \Pr(Y = i).$$

Now we write down the expression for $E[Y]$: [3]

$$E[Y] = \sum_{i=0}^{l-1} i \Pr(Y = i) + \sum_{t=l}^{\tau} \Pr(X_l = t)(\sum_{i=l}^{\tau} i \Pr(Y = i | X_l = t)). \tag{D.2}$$

Since $1 \geq p_l \geq p_{l+1} \geq \ldots \geq p_\tau > 0$, $E[Y]$ is less than the case where $p_l = p_{l+1} = \ldots = p_\tau$, although l remains the same. So it is sufficient to prove Theorem 1 for the case where $p_l = p_{l+1} = \ldots = p_\tau$. For $t \leq \tau$ if we set $p_l = p_{l+1} = \ldots = p_\tau$ we have

$$\sum_{i=l}^{\tau} i \Pr(Y = i | X_l = t) = E[Y | X_l = t] = l + (\tau - t) p_l. \tag{D.3}$$

Therefore, by using $(D.2)$ and $(D.3)$ we have

$$E[Y] = \sum_{i=0}^{l-1} i \Pr(Y = i) + \sum_{t=l}^{\tau} (l + p_l(\tau - t)) \Pr(X_l = t)$$

[3] Note that the summation on t here is from l to τ; since for $0 \leq t < \ell$, $\Pr(X_l = t) = 0$.

$$
\begin{aligned}
&= \sum_{i=0}^{l-1} i \Pr(Y = i) + (l + p_l\tau)(1 - \Pr(X_l > \tau)) \\
&\quad - p_l[\sum_{t=l}^{\infty} t \Pr(X_l = t) - \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t)] \\
&= \sum_{i=0}^{l-1} i \Pr(Y = i) + (l + p_l\tau) - (l + p_l\tau) \sum_{i=0}^{l-1} \Pr(Y = i) \\
&\quad - p_l \sum_{i=1}^{l} \frac{1}{p_i} + p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t) \\
&= \sum_{i=0}^{l-1} (i - l - p_l\tau) \Pr(Y = i) + (l + p_l(\tau - \sum_{i=1}^{l} \frac{1}{p_i})) + p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t) \\
&\overset{(a)}{<} \sum_{i=0}^{l-1} (i - l - p_l\tau) \Pr(Y = i) + l + 1 + p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t), \qquad \text{(D.4)}
\end{aligned}
$$

where the last inequality (a) follows from $\tau < \sum_{i=1}^{l+1} \frac{1}{p_i}$ and the assumption that $p_{l+1} = p_l$. Now, we only need to rewrite $p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t)$ in terms of Y. For $t > \tau$ we have

$$\Pr(X_l = t) = \sum_{i=0}^{l-1} \Pr(X_l = t | Y = i) \Pr(Y = i).$$

Therefore,

$$
\begin{aligned}
\sum_{t=\tau+1}^{\infty} t \Pr(X_l = t) &= \sum_{t=\tau+1}^{\infty} t (\sum_{i=0}^{l-1} \Pr(X_l = t | Y = i) \Pr(Y = i)) \\
&= \sum_{i=0}^{l-1} \Pr(Y = i) (\sum_{t=\tau+1}^{\infty} t \Pr(X_l = t | Y = i)).
\end{aligned}
$$

But due to memoryless property of geometric distribution, we know that

$$
\begin{aligned}
&\sum_{t=\tau+1}^{\infty} (t - \tau) \Pr(X_l = t | Y = i) \\
&= \sum_{t=\tau+1}^{\infty} (t - \tau) \Pr(\sum_{j=i+1}^{l} G_j = t - \tau) \\
&= \sum_{t=1}^{\infty} t \Pr(\sum_{j=i+1}^{l} G_j = t) = \sum_{j=i+1}^{l} \frac{1}{p_j}, \qquad \forall i \leq l - 1.
\end{aligned}
$$

Therefore, $\sum_{t=\tau+1}^{\infty} t\Pr(X_l = t|Y = i) = \tau + \sum_{j=i+1}^{l} \frac{1}{p_j}$. Hence,

$$p_l \sum_{t=\tau+1}^{\infty} t\Pr(X_l = t) = \sum_{i=0}^{l-1} \Pr(Y = i)(p_l\tau + \sum_{j=i+1}^{l} \frac{p_l}{p_j}). \tag{D.5}$$

Substituting (D.5) into (D.4) we get

$$\begin{aligned}
E[Y] < & \sum_{i=0}^{l-1}(i - l - p_l\tau)\Pr(Y = i) + l + 1 + \sum_{i=0}^{l-1} \Pr(Y = i)(p_l\tau + \sum_{j=i+1}^{l} \frac{p_l}{p_j}) \\
= & l + 1 + \sum_{i=0}^{l-1} \Pr(Y = i)(i - l - p_l\tau + p_l\tau + \sum_{j=i+1}^{l} \frac{p_l}{p_j}) \\
\leq & l + 1,
\end{aligned}$$

where the last inequality follows from the fact that $\forall j \in \{i+1, \ldots, l\} \quad p_l \leq p_j$.

D.4 Proof of Lemma 22

We will show that $E[Y] > l - 2\sqrt{l + \frac{1}{4}}$. It is sufficient to prove Lemma 22 for $q = l$; because for $q > l$, $E[Y]$ would only increase. On the other hand, q cannot be less than l according to the assumption $l = \max \hat{l} \quad s.t. \quad \sum_{i=1}^{\hat{l}} 1/p_i \leq \tau$. Therefore, from now on we suppose $q = l$. By our notation we have

$$\Pr(\sum_{j=1}^{i} G_j > \tau) = \sum_{j=0}^{i-1} \Pr(Y = j), \quad i = 1, 2, \ldots, l. \tag{D.6}$$

We now bound $l - E[Y]$ from above.

$$\begin{aligned}
l - E[Y] = l - \sum_{i=1}^{l} \Pr(Y \geq i) &= \sum_{i=1}^{l}(1 - \Pr(Y \geq i)) \\
&= \sum_{i=1}^{l} \Pr(Y < i) \overset{(a)}{=} \sum_{i=1}^{l} \Pr(\sum_{j=1}^{i} G_j > \tau) \\
&\overset{(b)}{\leq} \sum_{i=1}^{l} \Pr(\sum_{j=1}^{i} G_j > \sum_{j=1}^{l} \frac{1}{p_j})
\end{aligned}$$

$$\le 1 + \sum_{i=1}^{l-1} \Pr(|\sum_{j=1}^{i}(G_j - \frac{1}{p_j})| > \sum_{j=i+1}^{l} \frac{1}{p_j})$$

$$\overset{(c)}{\le} 1 + \sum_{i=1}^{l-1} \min(1, \frac{\mathrm{var}(\sum_{j=1}^{i} G_j)}{(\sum_{j=i+1}^{l} \frac{1}{p_j})^2})$$

$$\overset{(d)}{\le} 1 + \sum_{i=1}^{l-1} \min(1, \frac{\sum_{j=1}^{i} \frac{1}{p_j^2}}{(\sum_{j=i+1}^{l} \frac{1}{p_j})^2}),$$

where (a) follows from (D.6); (b) follows from $\sum_{i=1}^{l} 1/p_i \le \tau$; (c) follows from Chebyshev's inequality, where $\mathrm{var}(\sum_{j=1}^{i} G_j)$ is the variance of the random variable $\sum_{j=1}^{i} G_j$; and (d) follows due to independence of G_i's, which results in $\mathrm{var}(\sum_{j=1}^{i} G_j) = \sum_{j=1}^{i} \mathrm{var}(G_j) = \sum_{j=1}^{i} \frac{1-p_j}{p_j^2} < \sum_{j=1}^{i} \frac{1}{p_j^2}$, $i = 1, 2, \ldots, l$. But since $p_1 \ge p_2 \ge \ldots \ge p_l$, we have $\sum_{j=1}^{i} \frac{1}{p_j^2} \le \frac{i}{p_i^2}$ and $(\sum_{j=i+1}^{l} \frac{1}{p_j})^2 \ge (\frac{l-i}{p_i})^2$. Therefore,

$$l - E[Y] \le 1 + \sum_{i=1}^{l-1} \min(1, \frac{\frac{i}{p_i^2}}{\frac{(l-i)^2}{p_i^2}}) = 1 + \sum_{i=1}^{l-1} \min(1, \frac{i}{(l-i)^2}). \tag{D.7}$$

Hence, by (D.7) and applying Lemma 34 the proof of Lemma 22 will be complete.

Lemma 34. *Assume $l \in \mathbb{N}$, and $l > 1$. Then,* $1 + \sum_{i=1}^{l-1} \min(1, \frac{i}{(l-i)^2}) < 2\sqrt{l + \frac{1}{4}}$.

Proof. For $l < 18$ the statement of the Lemma can be verified numerically. Therefore, suppose that $l \ge 18$. Let $f(i) \triangleq \frac{i}{(l-i)^2}$, for $i \in \mathbb{N}, 1 \le i \le l-1$; and consider the following three observations regarding the function $f(.)$:

1. $f(i)$ increases as i increases for $i \in \mathbb{N}, 1 \le i \le l-1$.
2. $f(1) = \frac{1}{(l-1)^2} < 1$.
3. $f(l-1) = \frac{l-1}{1} > 1$.

Therefore, $\exists m \in \mathbb{N}, 1 \le m < l-1$ such that

$$\frac{m}{(l-m)^2} \le 1 < \frac{m+1}{(l-(m+1))^2}. \tag{D.8}$$

Note that $m \neq l-1$, because $\frac{l-1}{(l-(l-1))^2} > 1$. We rewrite the inequalities in (D.8) as

$$l - \sqrt{l+\frac{1}{4}} - \frac{1}{2} < m \leq l - \sqrt{l+\frac{1}{4}} + \frac{1}{2}. \tag{D.9}$$

In addition,

$$1 + \sum_{i=1}^{l-1} \min(1, \frac{i}{(l-i)^2}) = 1 + \sum_{i=1}^{m} \min(1, \frac{i}{(l-i)^2}) + \sum_{i=m+1}^{l-1} \min(1, \frac{i}{(l-i)^2}).$$

But from (D.8) and the fact that $f(i) = \frac{i}{(l-i)^2}$ increases by increase of i, we have

$$\begin{aligned}
&1 + \sum_{i=1}^{l-1} \min(1, \frac{i}{(l-i)^2}) = 1 + \sum_{i=1}^{m} \frac{i}{(l-i)^2} + (l-1-m) \\
&= l - m + \sum_{j=l-m}^{l-1} \frac{l-j}{j^2} < l - m + \sum_{j=l-m}^{l-1} \frac{l-j}{j(j-1)} \\
&= l - m + \sum_{j=l-m}^{l-1} (\frac{l-j}{j-1} - \frac{l-j}{j}) \\
&= l - m + \frac{m}{l-m-1} - \sum_{j=l-m}^{l-1} \frac{1}{j} \\
&\overset{(a)}{<} l - m + \frac{m}{l-m-1} - \frac{m}{l - \frac{m+1}{2}} \\
&= l - m + \frac{m(m+1)}{(l-m-1)(l-m+l-1)} \\
&\overset{(b)}{<} (\sqrt{l+\frac{1}{4}} + \frac{1}{2}) + \frac{(l - \sqrt{l+\frac{1}{4}} + \frac{1}{2})(l - \sqrt{l+\frac{1}{4}} + \frac{3}{2})}{(\sqrt{l+\frac{1}{4}} - \frac{3}{2})(l + \sqrt{l+\frac{1}{4}} - \frac{3}{2})} \\
&= (\sqrt{l+\frac{1}{4}} + \frac{1}{2}) \\
&\quad + (\sqrt{l+\frac{1}{4}} - \frac{3}{2} + \frac{5l - 9\sqrt{l+\frac{1}{4}} + \frac{11}{2}}{(\sqrt{l+\frac{1}{4}} - \frac{3}{2})(l + \sqrt{l+\frac{1}{4}} - \frac{3}{2})}) \\
&= 2\sqrt{l+\frac{1}{4}} + \frac{-l\sqrt{l+\frac{1}{4}} + \frac{11}{2}l - 6\sqrt{l+\frac{1}{4}} + 3}{(\sqrt{l+\frac{1}{4}} - \frac{3}{2})(l + \sqrt{l+\frac{1}{4}} - \frac{3}{2})},
\end{aligned} \tag{D.10}$$

where (a) follows from the Cauchy-Schwarz inequality; and (b) follows from (D.9). For $l \geq 18$ the term $\frac{-l\sqrt{l+\frac{1}{4}} + \frac{11}{2}l - 6\sqrt{l+\frac{1}{4}} + 3}{(\sqrt{l+\frac{1}{4}} - \frac{3}{2})(l + \sqrt{l+\frac{1}{4}} - \frac{3}{2})}$ in (D.10) is less than zero. Therefore, the statement of Lemma 34 is true for all $l > 1, l \in \mathbb{N}$. □

D.5 Proof of Corollary 3

Let $\vec{\Pi}^*$ denote the partition (assignment) chosen by the optimal greedy static scheduling policy $\eta^*_{\text{g-static}}$. Therefore, we have $\|\vec{R}(\eta^*_{\text{g-static}})\|_1 = C_{\mathsf{T}^3}$. Furthermore, consider an assignment, denoted by $\vec{\Pi}_{det}$, which maximizes the objective function in (5.4). Let $\eta^{\text{det}}_{\text{g-static}}$ denote the greedy static scheduling policy which corresponds to $\vec{\Pi}_{det}$. Further, let $\|\vec{R}_{\text{det}}(\eta_{\text{static}})\|_1$ designate the maximum number of objects that can be packed in the RP in (5.4) when a static scheduling policy η_{static} is implemented. Therefore, $\|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1 = C_{\text{det}}$, since $\|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1$ is the value of the objective function in (5.4) when the assignment is dictated by $\eta^{\text{det}}_{\text{g-static}}$. The right part of the inequality in Corollary 3 in (5.9) is trivial since C_{T^3} is the optimal T^3 achievable under any scheduling policy. So we only need to prove the left part of the inequality in (5.9). Using a similar argument as the one in part B of Section 5.4, and by applying Cauchy-Schwarz inequality, we get

$$\|\vec{R}(\eta^{\text{det}}_{\text{g-static}})\|_1 \geq \|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1 - 2\sqrt{N(\|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1 + \frac{N}{4})}. \tag{D.11}$$

Now consider the function $g(.)$ defined as follows: $g(x) \triangleq x - 2\sqrt{(N(x+\frac{N}{4}))}, \quad x \in \mathbb{R}$. So, $g(x)$ is strictly increasing for $x > \frac{3N}{4}$. On the other hand, we know that

$$\begin{aligned} C_{\text{det}} = \|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1 &\geq \|\vec{R}_{\text{det}}(\eta^*_{\text{g-static}})\|_1 \\ &\geq \|\vec{R}(\eta^*_{\text{g-static}})\|_1 - N = C_{\mathsf{T}^3} - N, \end{aligned} \tag{D.12}$$

where the right inequality follows from Theorem 8. By $g(x)$ being an increasing function of x and (D.12) we get

$$\|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1 - 2\sqrt{N(\|\vec{R}_{\text{det}}(\eta^{\text{det}}_{\text{g-static}})\|_1 + \frac{N}{4})} \geq C_{\mathsf{T}^3} - N - 2\sqrt{N(C_{\mathsf{T}^3} - \frac{3N}{4})}. \tag{D.13}$$

Hence, by (D.11) and (D.13) we get $\|\vec{R}(\eta^{\text{det}}_{\text{g-static}})\|_1 \geq C_{\mathsf{T}^3} - N - 2\sqrt{N(C_{\mathsf{T}^3} - \frac{3N}{4})}$.

D.6 Proof of Theorem 11

By the same argument as in proof of Lemma 20, $C_{w\text{-}\mathsf{T}^3}$ can be achieved by a static scheduling policy. Therefore, by LLN, to achieve $C_{w\text{-}\mathsf{T}^3}$, it is sufficient to find the assignment and ordering which provide the highest expected weighted delivery for one interval. First, we show that for a given assignment $\vec{\Pi} = [\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_N]$ the optimal ordering of the packets of clients assigned to AP_i is according to the order of $\omega_j p_{ij}$, $j \in \mathcal{I}_i$. To do so, it is sufficient to prove that for any given order of the clients if we swap two adjacent clients such that the client with higher $\omega_j p_j$ is prioritized higher, then the expected weighted delivery will be no less than before swapping. The following lemma formally states this fact.

Lemma 35. *Let $\tau, q \in \mathbb{N}$, and $\omega_1, \omega_2, \ldots, \omega_q \in \mathbb{R}$. Also, for some $d \in \{1, 2, \ldots, q-1\}$, let $\omega'_i = \omega_i$, for $1 \leq i < d$ and $d+1 < i \leq q$; and $\omega'_d = \omega_{d+1}$, $\omega'_{d+1} = \omega_d$. Further, let $G_1, G_2, \ldots, G_q$ be independent geometric random variables with parameters $p_1, p_2, \ldots, p_q$, respectively. Suppose that $\omega_d p_d < \omega_{d+1} p_{d+1}$. In addition, let $G'_1, G'_2, \ldots, G'_q$ be independent geometric random variables, independent of G_i's, with parameters $p_1, p_2, \ldots, p_{d-1}, p_{d+1}, p_d, p_{d+2}, \ldots, p_q$, respectively. Then,*

$$\sum_{i=1}^{q} \omega_i \Pr(\sum_{j=1}^{i} G_j \leq \tau) \leq \sum_{i=1}^{q} \omega'_i \Pr(\sum_{j=1}^{i} G'_j \leq \tau).$$

Proof. Let $A = \sum_{i=1}^{q} \omega_i \Pr(\sum_{j=1}^{i} G_j \leq \tau)$, and $B = \sum_{i=1}^{q} \omega'_i \Pr(\sum_{j=1}^{i} G'_j \leq \tau)$. Then,

$$\begin{aligned} B - A = &\sum_{i=1}^{d-1} \omega'_i \Pr(\sum_{j=1}^{i} G'_j \leq \tau) + \omega'_d \Pr(\sum_{j=1}^{d} G'_j \leq \tau) + \omega'_{d+1} \Pr(\sum_{j=1}^{d+1} G'_j \leq \tau) \\ &+ \sum_{i=d+2}^{q} \omega'_i \Pr(\sum_{j=1}^{i} G'_j \leq \tau) - \sum_{i=1}^{d-1} \omega_i \Pr(\sum_{j=1}^{i} G_j \leq \tau) - \omega_d \Pr(\sum_{j=1}^{d} G_j \leq \tau) \\ &- \omega_{d+1} \Pr(\sum_{j=1}^{d+1} G_j \leq \tau) - \sum_{i=d+2}^{q} \omega_i \Pr(\sum_{j=1}^{i} G_j \leq \tau) \end{aligned}$$

$$
\begin{aligned}
=&\omega'_d \Pr(\sum_{j=1}^{d} G'_j \le \tau) + \omega'_{d+1} \Pr(\sum_{j=1}^{d+1} G'_j \le \tau) - \omega_d \Pr(\sum_{j=1}^{d} G_j \le \tau) \\
&- \omega_{d+1} \Pr(\sum_{j=1}^{d+1} G_j \le \tau) \\
=&\sum_{t=1}^{\tau} \Pr(\sum_{j=1}^{d-1} G'_j = t)[\omega'_d \Pr(G'_d \le \tau - t) + \omega'_{d+1} \Pr(G'_d + G'_{d+1} \le \tau - t)] \\
&- \sum_{t=1}^{\tau} \Pr(\sum_{j=1}^{d-1} G_j = t)[\omega_d \Pr(G_d \le \tau - t) + \omega_{d+1} \Pr(G_d + G_{d+1} \le \tau - t)] \\
=&\sum_{t=1}^{\tau} \Pr(\sum_{j=1}^{d-1} G_j = t)[\omega'_d \Pr(G'_d \le \tau - t) + \omega'_{d+1} \Pr(G'_d + G'_{d+1} \le \tau - t) \\
&- \omega_d \Pr(G_d \le \tau - t) - \omega_{d+1} \Pr(G_d + G_{d+1} \le \tau - t)].
\end{aligned}
$$

Therefore, it is sufficient to show that for all $t \in \mathbb{N}$,

$$
\omega'_d \Pr(G'_d \le t) + \omega'_{d+1} \Pr(G'_d + G'_{d+1} \le t) - \omega_d \Pr(G_d \le t) - \omega_{d+1} \Pr(G_d + G_{d+1} \le t) \ge 0.
$$

Note that

- $\omega'_d = \omega_{d+1}$, and $\omega'_{d+1} = \omega_d$.
- $\Pr(G'_d \le t) = 1 - (1 - p_{d+1})^t$, and $\Pr(G_d \le t) = 1 - (1 - p_d)^t$.
- $\Pr(G_d + G_{d+1} \le t) = \Pr(G'_d + G'_{d+1} \le t) = 1 - \frac{p_d(1-p_{d+1})^t - p_{d+1}(1-p_d)^t}{p_d - p_{d+1}}$.

Therefore,

$$
\omega'_d \Pr(G'_d \le t) + \omega'_{d+1} \Pr(G'_d + G'_{d+1} \le t) - \omega_d \Pr(G_d \le t) - \omega_{d+1} \Pr(G_d + G_{d+1} \le t) \tag{D.14}
$$

$$
= (\omega_{d+1} p_{d+1} - \omega_d p_d)(\frac{(1 - p_{d+1})^t - (1 - p_d)^t}{p_d - p_{d+1}}) > 0, \quad t \in \mathbb{N}, \tag{D.15}
$$

where the inequality follows from the assumption that $\omega_{d+1} p_{d+1} - \omega_d p_d > 0$. □

D.6.1 Proof of $C_{w\text{-}\mathsf{T}^3} < C_{w\text{-}\mathbf{det}} + N\omega_{max}$

We follow the same line of proof as in Section 5.4. Since $C_{w\text{-}\mathsf{T}^3}$ can be achieved using a static scheduling policy which uses ordering according to $\omega_j p_j$'s, it is sufficient to show that for any static scheduling policy $\eta_{\text{wg-static}}$ which uses its corresponding optimal ordering we have $w\text{-}\mathsf{T}^3(\eta_{\text{wg-static}}) < C_{w\text{-det}} + N\omega_{max}$. Suppose an arbitrary static scheduling policy $\eta_{\text{wg-static}}$ with the corresponding partition $\vec{\Pi}_{\text{wg-static}} = [\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_N]$, which uses the optimal ordering is implemented. By (5.20) we know that $w\text{-}\mathsf{T}^3(\eta_{\text{wg-static}}) = \sum_{j=1}^{M} \omega_j R_j(\eta_{\text{wg-static}})$. On the other hand for $j \in [1:M]$, by (5.1) we have $R_j(\eta_{\text{wg-static}}) = \limsup_{r\to\infty} \frac{\sum_{k=1}^{r} N_j(k,\eta_{\text{wg-static}})}{r}$. For $i \in [1:N]$ define $Y_i \triangleq \sum_{j\in\mathcal{I}_i} N_j(1, \eta_{\text{wg-static}})$ and $q_i \triangleq |\mathcal{I}_i|$. Denote the enumeration of clients assigned to AP_i by $\{\mathcal{I}_i(1), \mathcal{I}_i(2), \ldots, \mathcal{I}_i(q_i)\}$, where the enumeration is according to the optimal ordering for the weighted case. Since a static scheduling policy is implemented and channels are i.i.d over time, by LLN we have

$$R_{\mathcal{I}_i(j)}(\eta_{\text{wg-static}}) = \limsup_{r\to\infty} \frac{\sum_{k=1}^{r} N_{\mathcal{I}_i(j)}(k, \eta_{\text{wg-static}})}{r} = \Pr(Y_i \geq j), \quad 1 \leq j \leq q_i, \quad 1 \leq i \leq N.$$

Therefore, it is easy to see that $w\text{-}\mathsf{T}^3(\eta_{\text{wg-static}}) = \sum_{i=1}^{N} \sum_{j=1}^{q_i} (\sum_{k=1}^{j} \omega_{\mathcal{I}_i(k)}) \Pr(Y_i = j)$. Let G_{ij} be a geometric random variable with parameter p_{ij}, $i \in [1:N]$, $j \in [1:M]$. Then, for $i \in [1:N]$, $1 \leq k \leq q_i$, $Y_i = \max \quad k \quad s.t. \quad \sum_{j=1}^{k} G_{i\mathcal{I}_i(j)} \leq \tau$, since $\eta_{\text{wg-static}}$ persistently sends a packet until it is delivered, or the interval is over. The following lemma, which is the generalized version of Lemma 21, relates l_i and ω_j's to Y_i.

Lemma 36. *Let $1 \leq \omega_1, \omega_2, \ldots, \omega_q \leq \omega_{max}$ for some $\omega_{max} \in \mathbb{R}$. Also, let $\tau \in \mathbb{N}$ and $G_1, G_2, \ldots, G_q$ be independent geometric random variables with parameters $p_1, p_2, \ldots, p_q$ respectively, such that $\omega_1 p_1 \geq \omega_2 p_2 \geq \ldots \geq \omega_q p_q \geq 0$. Also define $l \triangleq \max \hat{l} \quad s.t. \quad \sum_{i=1}^{\hat{l}} 1/p_i \leq \tau$, and $Y \triangleq \max i \quad s.t. \quad \sum_{j=1}^{i} G_j \leq \tau, \quad i \in \{1, 2, \ldots, q\}$. Then, we have $\sum_{i=1}^{q} (\sum_{j=1}^{i} \omega_j) \Pr(Y = i) < \sum_{j=1}^{l} \omega_j + \omega_{max}$.*

Proof. Suppose that $l > 0$ (for $l = 0$ the proof is straightforward). We have

$$\sum_{i=1}^{l} \frac{1}{p_i} \le \tau < \sum_{i=1}^{l+1} \frac{1}{p_i}. \tag{D.16}$$

Without loss of generality we can omit p_i's that are equal to zero and assume $0 < p_1, p_2, \ldots, p_q \le 1$. Furthermore, according to the same argument as in proof of Theorem 1, it is sufficient to prove the lemma for the case of $q = \tau$. Let $X_l = \sum_{i=1}^{l} G_i$, where $G_i = Geom(p_i)$. We have

$$\begin{aligned} \sum_{i=1}^{\tau}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i) &= \sum_{j=1}^{l-1}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i) \\ &+ \sum_{t=1}^{\tau} \Pr(X_l = t)(\sum_{i=l}^{\tau}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i | X_l = t)) \end{aligned}$$

However, since $\omega_{max} \ge \omega_l p_l \ge \omega_{l+1} p_{l+1} \ge \ldots \ge \omega_\tau p_\tau > 0$, $\sum_{i=1}^{\tau}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i)$ is less than the case where $\omega_l p_l = \omega_{l+1} p_{l+1} = \ldots = \omega_\tau p_\tau$. With a similar argument as in the proof of Theorem 1 we get

$$\begin{aligned} &\sum_{i=1}^{\tau}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i) \le \sum_{i=1}^{l-1}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i) + \sum_{t=1}^{\tau}(\sum_{j=1}^{l} \omega_j + \omega_l p_l(\tau - t)) \Pr(X_l = t) \\ &= \sum_{i=1}^{l-1}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i) + (\sum_{j=1}^{l} \omega_j + \omega_l p_l \tau)(1 - \Pr(X_l > \tau)) \\ &- \omega_l p_l[\sum_{t=1}^{\infty} t \Pr(X_l = t) - \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t)] = \sum_{i=1}^{l-1}(\sum_{j=1}^{i} \omega_j) \Pr(Y = i) + (\sum_{j=1}^{l} \omega_j + \omega_l p_l \tau) \\ &- (\sum_{j=1}^{l} \omega_j + \omega_l p_l \tau) \sum_{i=0}^{l-1} \Pr(Y = i) - \omega_l p_l \sum_{i=1}^{l} \frac{1}{p_i} + \omega_l p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t) \\ &= \sum_{i=0}^{l-1}(\sum_{j=1}^{i} \omega_j - \sum_{j=1}^{l} \omega_j - \omega_l p_l \tau) \Pr(Y = i) + (\sum_{j=1}^{l} \omega_j + \omega_l p_l(\tau - \sum_{i=1}^{l} \frac{1}{p_i})) \\ &+ \omega_l p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t) \\ &\overset{(a)}{<} \sum_{i=0}^{l-1}(\sum_{j=1}^{i} \omega_j - \sum_{j=1}^{l} \omega_j - \omega_l p_l \tau) \Pr(Y = i) + (\sum_{j=1}^{l} \omega_j + \omega_{l+1}) + \omega_l p_l \sum_{t=\tau+1}^{\infty} t \Pr(X_l = t) \end{aligned}$$

$$\stackrel{(b)}{=}\sum_{i=0}^{l-1}(-\sum_{j=i+1}^{l}\omega_j-\omega_l p_l\tau)\Pr(Y=i)+(\sum_{j=1}^{l}\omega_j+\omega_{l+1})+\sum_{i=0}^{l-1}\Pr(Y=i)(\omega_l p_l\tau+\sum_{j=i+1}^{l}\frac{\omega_l p_l}{p_j})$$

$$=(\sum_{j=1}^{l}\omega_j+\omega_{l+1})+\sum_{i=0}^{l-1}(\sum_{j=i+1}^{l}\frac{\omega_l p_l}{p_j}-\sum_{j=i+1}^{l}\omega_j)\Pr(Y=i)\stackrel{(c)}{\leq}\sum_{j=1}^{l}\omega_j+\omega_{l+1}\leq\sum_{j=1}^{l}\omega_j+\omega_{l+1}.$$

where (a) follows from $\tau-\sum_{i=1}^{l}\frac{1}{p_i}<\frac{1}{p_{l+1}}$ and $\omega_l p_l=\omega_{l+1}p_{l+1}$; (b) follows from (D.5); and (c) follows from the fact that $\forall j\in\{i+1,\ldots,l\}\quad \frac{\omega_l p_l}{p_j}\leq\omega_j$. □

Hence, by Lemma 36 we have

$$\begin{aligned}w\text{-}\mathsf{T}^3(\eta_{\text{wg-static}})&=\sum_{i=1}^{N}\sum_{j=1}^{q_i}(\sum_{k=1}^{j}\omega_{\mathcal{I}_i(k)})\Pr(Y_i=j)\stackrel{(a)}{<}\sum_{i=1}^{N}\sum_{j=1}^{l_i}\omega_{\mathcal{I}_i(j)}+N\omega_{max}.\\&\stackrel{(b)}{\leq}C_{w\text{-det}}+N\omega_{max},\end{aligned}$$

where (a) follows from Lemma 36; and (b) follows from the fact that $\sum_{i=1}^{N}\sum_{j=1}^{l_i}\omega_{\mathcal{I}_i(j)}$ is the value of the objective function in (5.22) for a feasible solution.

D.6.2 Proof of $C_{w\text{-det}}-2\omega_{max}\sqrt{N(C_{w\text{-det}}+\frac{N}{4})}<C_{w\text{-}\mathsf{T}^3}$

The proof of the lower bound is similar to the proof of lower bound in Theorem 8. Consider the assignment proposed by the solution to (5.22), where the clients which have not been assigned to any AP for transmission are assigned to AP's arbitrarily. Let $\vec{\Pi}^{\text{det}}_{\text{wg-static}}=[\mathcal{I}_1^{\text{det}},\mathcal{I}_2^{\text{det}},\ldots,\mathcal{I}_N^{\text{det}}]$ denote the resulting partition, and also let $\eta^{\text{det}}_{\text{wg-static}}$ denote the corresponding static scheduling policy which orders clients based on their channel success probabilities. Therefore, $w\text{-}\mathsf{T}^3(\eta^{\text{det}}_{\text{wg-static}})\leq C_{w\text{-}\mathsf{T}^3}$. So, it is sufficient to prove that $C_{w\text{-det}}-2\omega_{max}\sqrt{N(C_{w\text{-det}}+\frac{N}{4})}<w\text{-}\mathsf{T}^3(\eta^{\text{det}}_{\text{wg-static}})$.

For $i\in[1:N]$ let $W_i\triangleq\sum_{j\in\mathcal{I}_i^{\text{det}}}\omega_j N_j(1,\eta_{\text{wg-static}})$. Then, by LLN we have $w\text{-}\mathsf{T}^3(\eta^{\text{det}}_{\text{wg-static}})=\sum_{i=1}^{N}E[W_i^{\text{det}}]$. Therefore, it is sufficient to prove $C_{w\text{-det}}-$

$2\omega_{max}\sqrt{N(C_{w\text{-det}}+\frac{N}{4})} < \sum_{i=1}^{N} E[W_i^{\text{det}}]$. Define $q_i = |\mathcal{I}_i^{\text{det}}|$, and enumerate the clients assigned to AP_i by $\{\mathcal{I}_i^{\text{det}}(1), \mathcal{I}_i^{\text{det}}(2), \ldots, \mathcal{I}_i^{\text{det}}(q_i)\}$, where the enumeration is according to the channel success probabilities of different clients in $\mathcal{I}_i^{\text{det}}$. Further, let G_{ij} be a geometric random variable with parameter $p_{ij}, i \in [1:N], j \in [1:M]$. It is easy to see that for $k \le q_i, i \in [1:N]$, $W_i^{\text{det}} = \max \sum_{j=1}^{k} \omega_j \quad s.t. \quad \sum_{j=1}^{k} G_{i\mathcal{I}_i^{\text{det}}(j)} \le \tau, \quad i \in \{1,2,\ldots,N\}, k \le q_i$, since $\eta_{\text{g-static}}^{\text{det}}$ persistently sends a packet until it is delivered, or the interval is over. Also define $l_i^{\text{det}} \triangleq \max \quad \hat{l} \quad s.t. \quad \sum_{j=1}^{\hat{l}} 1/p_{i\mathcal{I}_i^{\text{det}}(j)} \le \tau, \quad \hat{l} \le q_i$. Then,

$$
\begin{aligned}
\sum_{i=1}^{N} E[W_i^{\text{det}}] &\overset{(a)}{>} \sum_{i=1}^{N}\sum_{j=1}^{l_i^{\text{det}}} \omega_{\mathcal{I}_i^{\text{det}}(j)} - 2\omega_{max}\sum_{i=1}^{N}\sqrt{l_i^{\text{det}}+\frac{1}{4}} \\
&\overset{(b)}{\ge} \sum_{i=1}^{N}\sum_{j=1}^{l_i^{\text{det}}} \omega_{\mathcal{I}_i^{\text{det}}(j)} - 2\omega_{max}\sqrt{N(\sum_{i=1}^{N} l_i^{\text{det}}+\frac{N}{4})} \\
&\ge \sum_{i=1}^{N}\sum_{j=1}^{l_i^{\text{det}}} \omega_{\mathcal{I}_i^{\text{det}}(j)} - 2\omega_{max}\sqrt{N(\sum_{i=1}^{N}\sum_{j=1}^{l_i^{\text{det}}} \omega_{\mathcal{I}_i^{\text{det}}(j)}+\frac{N}{4})} \\
&\overset{(c)}{=} C_{w\text{-det}} - 2\omega_{max}\sqrt{N(C_{w\text{-det}}+\frac{N}{4})},
\end{aligned}
$$

where (a) follows from Lemma 37; (b) follows from Cauchy-Schwarz inequality; and (c) follows from $\sum_{i=1}^{N}\sum_{j=1}^{l_i^{\text{det}}} \omega_{\mathcal{I}_i^{\text{det}}(j)} = C_{w\text{-det}}$. Hence, the left inequality of Theorem 11 is proved and the proof of Theorem 11 is complete.

Lemma 37. *Let $1 \le \omega_1, \omega_2, \ldots, \omega_q \le \omega_{max}$ for some $\omega_{max} \in \mathbb{R}$. Also, let $\tau \in \mathbb{N}$ and $G_1, G_2, \ldots, G_q$ be independent geometric random variables with parameters $p_1, p_2, \ldots, p_q$ respectively, such that $1 \ge p_1 \ge p_2 \ge \ldots \ge p_q \ge 0$. Also define $l \triangleq \max \hat{l} \quad s.t. \quad \sum_{i=1}^{\hat{l}} 1/p_i \le \tau$, and $Y \triangleq \max i \quad s.t. \quad \sum_{j=1}^{i} G_j \le \tau, \quad i \in \{1,2,\ldots,q\}$. Then, we have $\sum_{j=1}^{l}\omega_j - 2\omega_{max}\sqrt{l+\frac{1}{4}} < \sum_{i=1}^{q}(\sum_{j=1}^{i}\omega_j)\Pr(Y=i)$.*

Proof. With the same argument as in Theorem 1, it is sufficient to assume $q = l$.

The proof is very similar to the proof of lower bound in Theorem 1:

$$\begin{aligned}
&\sum_{i=1}^{l}\omega_i - \sum_{i=1}^{l}(\sum_{j=1}^{i}\omega_j)\Pr(Y=i) = \sum_{i=1}^{l}\omega_i - \sum_{i=1}^{l}\omega_i(\sum_{j=i}^{l}\Pr(Y=j)) \\
&= \sum_{i=1}^{l}\omega_i \Pr(G_1+G_2+\ldots G_i > \tau) \\
&\le \sum_{i=1}^{l-1}\omega_i \Pr(|\sum_{j=1}^{i}(G_j - \frac{1}{p_j})| > \sum_{j=i+1}^{l}\frac{1}{p_j}) + \omega_l \overset{(a)}{\le} \sum_{i=1}^{l-1}\omega_i \min(1, \frac{var(\sum_{j=1}^{i} G_j)}{(\sum_{j=i+1}^{l}\frac{1}{p_j})^2}) + \omega_l \\
&\le \sum_{i=1}^{l-1}\omega_i \min(1, \frac{\sum_{j=1}^{i}\frac{1}{p_j^2}}{(\sum_{j=i+1}^{l}\frac{1}{p_j})^2}) + \omega_l \le \omega_l + \sum_{i=1}^{l-1}\omega_i \min(1, \frac{i}{(l-i)^2}) \\
&\le \omega_{max}(1 + \sum_{i=1}^{l-1}\min(1, \frac{i}{(l-i)^2})) \overset{(b)}{\le} 2\omega_{max}\sqrt{l+\frac{1}{4}},
\end{aligned}$$

where (a) follows from Chebyshev's inequality; and (b) follows from Lemma 34. ◻

BIBLIOGRAPHY

[1] Cisco Visual Networking Index: Forecast and Methodology, 2010-2015. *available at www.cisco.com*, 2011.

[2] N. Bambos A. Dua, C.W. Chan and J. Apostolopoulos. *IEEE Trans. on Wireless Communications, Title = Channel, deadline, and distortion* (CD^2) *aware scheduling for video streams over wireless, Volume = 9, Number = 3, Year = 2010*.

[3] Akbar Ghasemi Abdoli, Mohammad Javad and Amir K. Khandani. Full-duplex transmitter cooperation, feedback, and the degrees of freedom of SISO Gaussian interference and X channels. *IEEE International Symposium on Information Theory (ISIT)*, 2012.

[4] J. Abdoli and A. S. Avestimehr. Layered interference networks with delayed CSI: DoF scaling with distributed transmitters. *IEEE Transactions on Information Theory*, 2013.

[5] J. Abdoli and A. S. Avestimehr. On degrees of freedom scaling in layered interference networks with delayed CSI. *IEEE International Symposium on Information Theory*, 2013.

[6] M.J. Abdoli, A. Ghasemi, and A.K. Khandani. On the degrees of freedom of K-user SISO interference and X channels with delayed CSIT. *arXiv:1109.4314*, 2011.

[7] M.J. Abdoli, A. Ghasemi, and A.K. Khandani. On the Degrees of Freedom of K-User SISO Interference and X Channels with Delayed CSIT. *arXiv:1109.4314*, 2011.

[8] M. Agarwal and A. Puri. Base station scheduling of requests with fixed deadlines. *In Proc. of IEEE INFOCOM*, 2:487 – 496, 2002.

[9] SaiDhiraj Amuru, R. Tandon, and S. Shamai. On the degrees-of-freedom of the 3-user MISO broadcast channel with hybrid CSIT. *arXiv preprint arXiv:1402.4729*, 2014.

[10] et al. Avestimehr, Salman. Video delivery over wireless networks: exploiting network heterogeneity and content commonality. *Intel Technology Journal*, 19(1), 2015.

[11] G. Bagherikaram, A.S. Motahari, and A.K. Khandani. On the secure Degrees-of-Freedom of the multiple-access-channel. *arXiv preprint arXiv:1003.0729*, 2010.

[12] J.F. Benders and J.A.E.E. van Nunen. A property of assignment type mixed integer linear programming problems. *O.R. Letters*, (2):47–52, 1982.

[13] G. Bresler, D. Cartwright, and D. N. Tse. Interference alignment for the MIMO interference channel. *arXiv:1303.5678*, 2013.

[14] G. Bresler and D. Tse. The two-user Gaussian interference channel: a deterministic view. *European Transactions in Telecommunications*, 19(4):333354, 2008.

[15] V. R. Cadambe and S. A. Jafar. Degrees of freedom of wireless X networks. *IEEE International Symposium on Information Theory*, 2008.

[16] Viveck R. Cadambe and Syed A. Jafar. Interference alignment and the degree of freedom for the K user interference channel. *IEEE Transactions on Information Theory*, 54(8):3425–3441, August 2008.

[17] C. Chekuri and S. Khanna. A PTAS for the Multiple Knapsack Problem. *SIAM Journal on Computing*, 2005.

[18] I. Csiszár and J. Korner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, 1978.

[19] Imre Csiszr and Prakash Narayan. Secrecy capacities for multiple terminals. *IEEE Transactions on Information Theory*, 50(12):3047–3061, 2004.

[20] Arash Gholami Davoodi and Syed A. Jafar. Aligned Image Sets under Channel Uncertainty: Settling a Conjecture by Lapidoth, Shamai and Wigger on the Collapse of Degrees of Freedom under Finite Precision CSIT. *arXiv preprint arXiv:1403.1541*, 2014.

[21] Arash Gholami Davoodi and Syed A. Jafar. Aligned image sets under channel uncertainty: settling a conjecture by lapidoth, shamai and wigger on the collapse of degrees of freedom under finite precision CSIT. *arXiv preprint arXiv:1403.1541*, 2014.

[22] H. Federer. Geometric measure theory. *Reprint of the 1969 Edition. Springer*, 1996.

[23] Abbas El Gamal. The feedback capacity of degraded broadcast channels. *IEEE Transactions on Information Theory*, 24(3):379–381, May 1978.

[24] L. Georgiadis and L. Tassiulas. Broadcast erasure channel with feedback - capacity and algorithms. *Worshop on Network Coding, Theory, and Applications*, pages 54–61, 2009.

[25] A. Ghasemi, A.S. Motahari, and A.K. Khandani. On the degrees of freedom of *X* channel with delayed CSIT. *Proc. of IEEE ISIT*, 2011.

[26] T. Gou and S.A. Jafar. On the secure degrees of freedom of wireless X networks. *Proc. of Allerton Conference on Communicaiton, Control, and Computing*, 2008.

[27] X. He, A. Khisti, and A. Yener. MIMO broadcast channel with arbitrarily varying eavesdropper channel: secrecy degrees of freedom. *Proc. of IEEE GLOBECOM*, 2011.

[28] X. He and A. Yener. MIMO wiretap channels with arbitrarily varying eavesdropper channel states. *arXiv preprint arXiv:1007.4801*, 2010.

[29] I-H. Hou and P.R. Kumar. Admission control and scheduling for QoS guarantees for variable-bit-rate applications on wireless channels. *In Proc. of ACM MobiHoc*, 2009.

[30] I-H. Hou and P.R. Kumar. Scheduling heterogeneous real-time traffic over fading wireless channels. *In Proc. of IEEE INFOCOM*, 2010.

[31] I-Hong Hou and P.R. Kumar. Scheduling periodic real-time tasks with heterogeneous reward requirements. *In Proc. of RTSS*, 2011.

[32] Chiachi Huang and Syed Ali Jafar. Degrees of freedom of the MIMO interference channel with cooperation and cognition. *IEEE Transactions on Information Theory*, 55(9):4211–4220, 2009.

[33] S. Chakraborty I-H. Hou, A. Truong and P.R. Kumar. Optimality of periodwise static priority policies in real-time communications. *In Proc. of CDC*, 2011.

[34] V. Borkar I-H. Hou and P.R. Kumar. A theory of QoS for wireless. *In Proc. of IEEE INFOCOM*, 2009.

[35] S. A. Jafar and S. Shamai. Degrees of freedom region of the MIMO *X* channel. *IEEE Transactions on Information Theory*, 54(1):151–170, January 2008.

[36] K Jain. A factor 2 approximation algorithm for the generalized Steiner network problem. *Combinatorica, Springer*, 2001.

[37] Urbashi Mitra Jindal, Nihar and Andrea Goldsmith. Capacity of ad-hoc networks with node cooperation. *Proc. of IEEE ISIT*, 2004.

[38] David TH Kao and Amir Salman Avestimehr. Linear degrees of freedom of the MIMO X-channel with delayed CSIT. *Proc. of IEEE ISIT*, 2014.

[39] A. Khisti. Interference alignment for the multiantenna compound wiretap channel. *IEEE Transactions on Information Theory*, 57(5):2976–2993, 2011.

[40] A. Khisti and G.W. Wornell. Secure transmission with multiple antennas I: The MISOME wiretap channel. *IEEE Transactions on Information Theory*, 56(7):3088–3104, 2010.

[41] A. Khisti and D. Zhang. Artificial-noise alignment for secure multicast using multiple antennas. *arXiv preprint, arXiv:1211.4649*, 2012.

[42] M. Kobayashi, P. Piantanida, S. Yang, and S. Shamai. On the secrecy degrees of freedom of the multiantenna block fading wiretap channels. *IEEE Transactions on Information Forensics and Security*, 6(3):703–711, 2011.

[43] Onur Ozan Koyluoglu, Hesham El Gamal, Lifeng Lai, and H Vincent Poor. On the secure degrees of freedom in the K-user Gaussian interference channel. In *Proc. of IEEE ISIT*, pages 384–388. IEEE, 2008.

[44] V. S. Mirrokni L. K. Fleischer, M. X. Goemans and M. Sviridenko. (Almost) tight approximation algorithms for maximizing general assignment problems. *Symposium on Discrete Algorithms (SODA)*, 2006.

[45] S. Lashgari and A. S. Avestimehr. Blind MIMO wiretap channel with delayed CSIT. *IEEE Globecom second workshop on trusted communications with Physical Layer Security (TCPLS20014).*

[46] S. Lashgari and A. S. Avestimehr. Blind MIMO Wiretap Channel with Delayed CSIT. *IEEE Globecom, Second workshop on trusted communications with Physical Layer Security*, 2014.

[47] S. Lashgari and A. S. Avestimehr. Blind Wiretap Channel with Delayed CSIT. *Proc. of IEEE ISIT*, 2014.

[48] S. Lashgari and A. S. Avestimehr. Blind wiretap channel with delayed CSIT. *submitted to IEEE International Symposium on Information Theory*, 2014.

[49] S. Lashgari and A. S. Avestimehr. Blind wiretap channel with delayed CSIT. *available on ArXiv, arXiv:1405.0521*, 2014.

[50] S. Lashgari and A. S. Avestimehr. Transmitter cooperation in interference channel with delayed CSIT. *Proc. of Allerton Conference on Communicaiton, Control, and Computing*, 2014.

[51] S. Lashgari, A. S. Avestimehr, and C. Suh. A rank ratio inequality and the linear degrees of freedom of X-channel with delayed CSIT. *Allerton Conference on Communicaiton, Control, and Computing*, 2013.

[52] S. Lashgari, A. S. Avestimehr, and C. Suh. Linear Degrees of Freedom of the X-Channel with Delayed CSIT. *IEEE Transactions on Information Theory*, 2014.

[53] S. Lashgari and A.S. Avestimehr. Approximating the timely throughput of heterogeneous wireless networks. *Proc. of IEEE ISIT*, 2012.

[54] S. Lashgari and A.S. Avestimehr. Timely Throughput Optimization of Heterogeneous Wireless Networks. *IEEE COMSOC MMTC E-Letter*, 8(5), 2013.

[55] Sina Lashgari and Amir Salman Avestimehr. Timely throughput of heterogeneous wireless networks: Fundamental limits and algorithms. *IEEE Transactions on Information Theory*, 59(12):8414–8433, 2013.

[56] Sina Lashgari and Amir Salman Avestimehr. Blind MIMOME wiretap channel with delayed CSIT. *submitted to IEEE Transactions on Information Theory*, 2015.

[57] Sina Lashgari, Ravi Tandon, and Salman Avestimehr. A general outer bound for MISO broadcast channel with heterogeneous CSIT. *IEEE ISIT 2015*, 2015.

[58] Sina Lashgari, Ravi Tandon, and Salman Avestimehr. MISO Broadcast Channel with Hybrid CSIT: Beyond Two Users. *arXiv preprint arXiv:1504.04615*, 2015.

[59] Sina Lashgari, Ravi Tandon, and Salman Avestimehr. Three-user MISO broadcast channel: how much can CSIT heterogeneity help? *IEEE ICC 2015*, 2015.

[60] S. Leung-Yan-Cheong and M. Hellman. The Gaussian wire-tap channel. *IEEE Transactions on Information Theory*, 24(4):451–456, 1978.

[61] Yingbin Liang and H. Vincent Poor. Information theoretic security. *Foundations and Trends in Communications and Information Theory*, 2009.

[62] T. Liu and P. Viswanath. An extremal inequality motivated by multiterminal information-theoretic problems. *IEEE Transactions on Information Theory*, 53:1839–1851, 2007.

[63] L. Lovasz. Submodular functions and convexity. *Springer*, 1983.

[64] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani. Communication over MIMO *X* channels: interference alignment, decomposition, and performance analysis. *IEEE Transactions on Information Theory*, 54(8):3457–3470, August 2008.

[65] M. A. Maddah-Ali and D. N. Tse. Completely stale transmitter channel state information is still very useful. *IEEE Transactions on Information Theory*, 58(7):4418–4431, 2012.

[66] H. Maleki, S. A. Jafar, and S. Shamai. Retrospective interference alignment over interference networks. *IEEE Journal of Selected Topics in Signal Processing*, 6(3):228–240, 2012.

[67] Roy D. Yates Maric, Ivana and Gerhard Kramer. Capacity of interference channels with partial transmitter cooperation. *IEEE Transactions on Information Theory*, 2007.

[68] Ueli Maurer and Stefan Wolf. Information-theoretic key agreement: from weak to strong secrecy for free. *Advances in CryptologyEUROCRYPT 2000. Springer Berlin Heidelberg*, 2000.

[69] Kaniska Mohanty and Mahesh K. Varanasi. On the DoF region of the K-user MISO broadcast channel with hybrid CSIT. *arXiv preprint arXiv:1312.1309*, 2013.

[70] Kaniska Mohanty, Chinmay S. Vaze, and Mahesh K. Varanasi. The degrees

of freedom region for the MIMO interference channel with hybrid CSIT. *arXiv preprint*, 2012.

[71] Pritam Mukherjee, Ravi Tandon, and Sennur Ulukus. Secure degrees of freedom region of the two-user MISO broadcast channel with alternating CSIT. *arXiv preprint arXiv:1502.02647*, 2015.

[72] Pritam Mukherjee, Ravi Tandon, and Sennur Ulukus. Secure degrees of freedom region of the two-user MISO broadcast channel with alternating CSIT. *arXiv preprint arXiv:1502.02647*, 2015.

[73] M. J. Neely. Delay Analysis for Max Weight Opportunistic Scheduling in Wireless Systems. *IEEE Trans. on Automatic Control*, (9), 2009.

[74] M. J. Neely. Dynamic optimization and learning for renewal systems. *In Proc. of ASILOMAR conference on signals, systems, and computers*, (11), 2010.

[75] Borzoo Rassouli, Chenxi Hao, and Bruno Clerckx. DoF analysis of the K-user MISO broadcast channel with hybrid CSIT. *arXiv preprint*, 2015.

[76] S. Shakkottai and R. Srikant. Scheduling real-time traffic with deadlines over a wireless channel. *Wireless Networks*, 8(1), 2002.

[77] D. B. Shmoys and E. Tardos. An approximation algorithm for the generalized assignment problem. *Mathematical Programming*, 62:461474, 1993.

[78] I. Shomorony and A. S. Avestimehr. Degrees of freedom of two-hop wireless networks: everyone gets the entire cake. *arXiv:1210.2143*, 2013.

[79] P. Stanica. Good Lower and Upper Bounds On Binomial Coefficients. *Journal of Inequalities in Pure and Applied Mathematics*, 2001.

[80] R. Tandon, S. Ulukus, and K. Ramchandran. Secure source coding with a helper. *IEEE Transactions on Information Theory*, 2012.

[81] Ravi Tandon, Syed Ali Jafar, Shlomo Shamai, and H. V. Poor. On the synergistic benefits of alternating CSIT for the MISO broadcast channel. *IEEE Transactions on Information Theory*, 59(7):4106–4128, 2013.

[82] Ravi Tandon, M. Maddah-Ali, A. Tulino, H. V. Poor, and S. Shamai. On fading broadcast channels with partial channel state information at the transmitter. *Proc. of IEEE ISWCS*, 2012.

[83] M. A. Trick. A Linear relaxation heuristic for the generalized assignment problem. *Naval Research Logistics*, 1992.

[84] A. Vahid, M. Maddah-Ali, and A. S. Avestimehr. Capacity results for binary fading interference channels with delayed CSIT. *IEEE Transactions on Information Theory*, 60(10):6093–6130, 2014.

[85] C. S. Vaze and M. K. Varanasi. The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT. *IEEE Transactions on Information Theory*, 58(8):5354–5374, 2012.

[86] C. S. Vaze and M. K. Varanasi. The degrees of freedom region and intererence alignment for the MIMO inteference channel with delayed CSIT. *IEEE Transactions on Information Theory*, 58(7):4396–4417, 2012.

[87] Chinmay S. Vaze and Mahesh K. Varanasi. The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSIT. *Proc. of IEEE ISIT*, 2011.

[88] I-Hsiang Wang and David NC Tse. Interference mitigation through limited transmitter cooperation. *IEEE Transactions on Information Theory*, 57.5:2941–2965, 2011.

[89] Sriram Vishwanath Wu, Wei and Ari Arapostathis. Capacity of a class of cognitive radio channels: Interference channels with degraded message sets. *IEEE Transactions on Information Theory*, 53(11):4391–4399, 2007.

[90] A. D. Wyner. The wire-tap channel. *Bell System Technical Journal*, 54(8):1355–1387, 1975.

[91] J. Xie and S. Ulukus. Secure degrees of freedom of one-hop wireless networks. *arXiv preprint, arXiv:1209.5370*, 2012.

[92] J. Xie and S. Ulukus. Secure degrees of freedom of the Gaussian wiretap channel with helpers and no eavesdropper CSI: Blind cooperative jamming. *arXiv preprint, arXiv:1302.6570*, 2013.

[93] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi. Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT. *IEEE Transactions on Information Theory*, 59(1):315–328, 2012.

[94] Sheng Yang, Pablo Piantanida, Mari Kobayashi, and Shlomo Shamai. On

the secrecy degrees of freedom of multi-antenna wiretap channels with delayed CSIT. In *Proc. of IEEE ISIT*, pages 2866–2870. IEEE, 2011.

[95] Sheng Yang, Pablo Piantanida, Mari Kobayashi, and Shlomo Shamai. Secrecy degrees of freedom of MIMO broadcast channels with delayed CSIT. *IEEE Transactions on Information Theory*, 59(9):5244–5256, 2013.

[96] D.D. Yao. Dynamic scheduling via polymatroid optimization. *in* performance evaluation of complex systems: techniques and tools*, LNCS, Springer Verlag*, (2459):89–113, 2002.

[97] Abdellatif Zaidi, Hassan Awan, Shlomo Shamai, and Luc Vandendorpe. Secure degrees of freedom of MIMO X-channels with output feedback and delayed CSIT. *IEEE Transactions on Information Forensics and Security*, 8(10):1760–1774, 2013.

[98] Abdellatif Zaidi, Hassan Awan, Shlomo Shamai, and Luc Vandendorpe. Secure degrees of freedom of X-channel with output feedback and delayed CSIT. In *Proc. of IEEE ITW*, 2013.